THE DIFFERENTIAL CALCULUS

OXFORD
UNIVERSITY PRESS
AMEN HOUSE, E.C. 4
LONDON EDINBURGH GLASGOW
NEW YORK TORONTO MELBOURNE
CAPETOWN BOMBAY CALCUTTA
MADRAS SHANGHAI
HUMPHREY MILFORD
PUBLISHER TO THE
UNIVERSITY

THE DIFFERENTIAL CALCULUS

BY

THEODORE CHAUNDY
STUDENT AND TUTOR OF CHRIST CHURCH
OXFORD

OXFORD
AT THE CLARENDON PRESS
1935

PRINTED IN GREAT BRITAIN

Hopefully,

to D., C., D., A. AND D.,

who have considered no obstacle too great

to be overcome in the preparation

of this work

πάντα ῥεῖ (HERACLITUS)

PREFACE

A PREFACE gives the author his last chance of disarming critics, or, at least, of anticipating them; and, in taking somewhat rueful leave of *The Differential Calculus*, I realize that it departs far enough from the accepted practices of the day to need some sort of propitiative apologia.

My original purpose (so far as it can be recalled at this distance of time) was to write, for my own instruction, a text-book combining the English plan with continental rigour. The text-book has grown into something far more unwieldy, but many of the English chapter-headings remain. Thus, the book is not a *Cours d'Analyse*, nor even a Theory of Real Functions. It is primarily a 'Calculus', concerned with showing how to do certain computative jobs. But, since it is both prudent and pleasant to understand the workings of one's machinery, the theoretical aspects of the subject have also been carefully considered. There are, of course, two schools of thought—the formalist and the rigorist—and I may be thought to have fallen between two schools in giving each too much space to please the other.

The uses of the Calculus in Geometry, except here and there by way of example, and in Mechanics are not dealt with. Its applications are restricted to Pure Mathematics and to the field of the real variable: for the Theory of Functions of a Complex Variable is essentially a development of the Integral Calculus arising out of Cauchy's theorem.

There is an almost complete absence of bibliographical material. Most of the important discoveries in the Differential Calculus are by now historical, and I am no historian. Moreover, the form in which a theorem may have been first enunciated by its inventor is not always the form that fits best into a later development of the subject. I have always admired Euclid, as I first read him, for a wise suppression of his authorities and his easy flow from proposition to proposition. Him I have taken as my exemplar, adopting the Platonic† aphorism: ὅπῃ ἂν ὁ λόγος ὥσπερ πνεῦμα φέρῃ, ταύτῃ ἰτέον.‡ This imputing of the responsibility to the 'argument' is a delightful excuse for inconsequence, as Plato§ no doubt intended, and I may well be accused of sacrificing important topics in the Calculus to irrelevancies dragged in from other corners of Mathematics.

In scope, then, I have included anything that interested me and anything that I wanted to know, and have excluded nothing merely because

† Or Socratic. ‡ Plato, *Republic*, 394 d. § Or Socrates.

it could belong to an elementary treatment of the subject. The development aims, therefore, at being complete and self-contained, and it should comprise within itself the material of a definite text-book based on my own tutorial experience: the blue-pencilling may be left to others.

The examples are in large part original (but, of course, not necessarily new), or have been modified from existing sources. Many of them are difficult: I have excluded mere 'examination-stuff', keeping only what seemed to have some interest. A 'problem' falls short of a 'theorem' only in lacking, at the moment, importance enough to be built up into the authoritative structure of the subject, receiving a factitious interest by being posed as a conundrum. The 'worked examples' give the solutions of some of the more difficult or interesting examples.

I set no store by orthodoxy, but for the protection of the unwary I have tried to erect warning signs before the more dangerous deviations from the normal. In particular, *I* replacing the more usual *we* is self-accusative and to be read 'I do, but others don't'. Thus, the terms *vanishingly uniform convergence* and *umbral derivative* lack authority, and I have given my own interpretation of *differential*. In chapter XII the modern theory of Differential Geometry has been ignored. That theory is so coherent in notation and method that it seemed impossible to tamper with it without ruining its peculiar elegance.

I have purposely experimented with notation, because, like all invention, notation achieves perfection only by accumulated improvement, and because notation biases analysis as language biases thought. In this experimentation I have been generously encouraged by the Clarendon Press. I take this opportunity of expressing my great obligation to them: in particular, to the Delegates and their officers, without whose perseverance the writing of this book might never have come to an end; and to their technical staff for their tireless attention to detail and for the beauty of the printing.

My thanks are likewise due to my former pupils Dr. Eric G. Phillips and Mr. G. J. Whitrow for reading the complete proofs. Of many others to whom I am indebted I mention only Mr. I. O. Griffith, kindest of collaborators, with whom the book was originally planned, and Professor G. H. Hardy, most patient of advisers, but for whom the book would never have seen light.

And so I cast my pebble on the beach.

CHRIST CHURCH,
OXFORD

T. C.

CONTENTS

XIV. EXPANSIONS IN POWER-SERIES.

I

FUNCTIONAL DEPENDENCE

1. Theories of number

MATHEMATICAL ANALYSIS, in which the Differential Calculus occupies an important place, is concerned, in essence, with numbers and with relations between numbers. We thus do well to take a preliminary, brief survey of the various classes of number with which we shall have to deal. They represent successive stages in the development of the art (and later of the science) of counting and measuring. Counting, of itself, gives only the *positive integers*, which are the fundamental units of the science. But in terms of them we are able to define other categories of number necessary to the process of measurement.

Efficient measurement requires an arithmetical equivalent to every point of a chosen straight line, the 'scale of length'. On this scale the positive integers mark isolated and equidistant points.

We augment them in the first place by the *positive fractions*, which render possible a change of the unit of measurement. Historically, they are the most ancient of the categories of 'invented number', coming to us from the Greeks. We define them arithmetically by pairing the positive integers in all possible ways in a symbol 'm/n' (where the order of m, n matters) and prescribing rules for the application to these symbols of the fundamental arithmetical operations of addition, subtraction, multiplication, division. As we know, addition and multiplication can always be effected between the positive integers; we may say that they form a 'complete' system for such operations. But we know equally that the system is incomplete for subtraction or division. With the introduction of positive fractions, however, we have the system of positive rational number, complete for division, but not for subtraction. It provides arbitrarily many points between the unit points of the scale of length.

We render this system of number complete for subtraction by the introduction of *negative number*, assigning to every positive rational number p a negative counterpart '$-p$'. This innovation belongs to the classical days of European mathematics and was for long a fiercely contested novelty. The importance of negative number in geometry enters with the use of coordinates. It permits us to pass smoothly from quadrant to quadrant, without change of formula in crossing an axis; in polar coordinates it finds a place for each sense of angular rotation. Our scale may now extend arbitrarily far in either direction.

We have completed the system of *rational number*, but we are still very far from having provided an arithmetical equivalent for every point of a straight line. It is true that between any two points of the line we can find arbitrarily many points nameable by rational numbers. But the Pythagorean formula $a^2+b^2 = c^2$ first indicates the presence of gaps in the rational scale of length. For instance, the side and diagonal of a square are incommensurable in rational terms.

2. Irrational number

We remedy this deficiency by introducing the most modern of the categories of number, that of *irrational number*. Whereas fractional and negative numbers were defined by reference, at most, to a pair of numbers of a type already known, an irrational number depends for its definition on *all* the rational numbers, or at least on an infinite set of them.

If P is a point of the line unaccounted for in the rational enumeration, every rational point must lie to the right or to the left of it. P thus splits the rational points into two groups; and these groups depend uniquely on P. For, if Q is another such point, say to the right of P, then infinitely many rational points lie in the segment PQ, and these points are to the right of P and to the left of Q. The pair of groups for Q is therefore not identical with the pair of groups for P.

This bipartite grouping of the rational points carries with it, of course, a corresponding bipartite grouping of the rational numbers, and it is such a grouping that we mean when we speak of an 'irrational' number. In precise terms, an irrational number denotes a separation of *all* the rational numbers into *two* groups that *do not overlap*. In other words, (i) every rational number is in one or other group; (ii) not all the rational numbers are in one group; (iii) every number in one group is greater than every number in the other. I shall say that such a separation is (i) complete, (ii) substantial, i.e. non-trivial, (iii) clear-cut.

Since infinitely many rational numbers are involved in the definition of a single irrational number, it is clear that the definition must turn on a law or a test by which we can assign every rational number to its proper group.

Thus, for instance, we define the irrational $\sqrt{2}$ by the test '$p^2 \gtrless 2$?', which separates the rationals into two groups of the character prescribed. For

(i) it is *complete*, since no rational has a square equal to 2 and therefore the square of every rational must be greater than or less than 2;

(ii) it is *substantial*, since $1^2 < 2 < 2^2$ and therefore neither group contains all the rationals;

(iii) it is *clear-cut*, since, if $p^2 < 2 < q^2$, then $p < q$.

The system of number thus far extended constitutes the system of *real* number. It provides an arithmetical equivalent for every point of a straight line. Arithmetically it is complete for passing to the limit in any sequence. We can, in fact, construct sequences of rational numbers that converge in the field of real number, but do not converge to a limit in the field of rational number. Such, for instance, is the sequence of successive convergents to any infinite continued fraction: in particular, the sequence $1, \frac{3}{2}, \frac{7}{5}, \frac{17}{12}, \frac{41}{29}, \ldots$, which converges in the real field to $\sqrt{2}$.

This system of number is still incomplete for the solution of algebraic equations. We can complete it, we know, by pairing all real numbers in a form $a+ib$, where i is a new symbol of special meaning. The implications of such an extension of the notion of number are algebraic rather than arithmetical. Moreover, the theory of differentiation in respect of such complex numbers is a special and distinct branch of mathematical study. In this book the field of thought is essentially the field of *real number*; we shall depart from that field only for occasional simplicity of formal statement, never for serious consideration of the problems it presents.

3. Variables

In the preceding discussion, geometrical language has been freely used, despite the fact that the subject-matter is purely analytical, and that every step in the discussion could, if desired, have been made in purely analytical terms. I shall continue to use this freedom of geometrical idiom without prejudicing the intrinsically analytical character of the work.

The cautious reader should always find it possible (if he so desire) to restate the argument in a form divested of spatial ideas. This reference to geometry, apart from the gain in brevity and vividness, is also reasonable, since much of the development of the Calculus has been in response to the demands of the geometer or the physicist.

The Calculus rightly begins with the notion of a *variable*. We can best illustrate it by the geometrical picture of a *moving point* that travels over the scale of length. It picks out, as coordinates of the points through which it passes, certain from among the real numbers. We may link these numbers into a unity by ascribing them as *values*

of the variable. In Mathematical Physics the notion is especially applicable, for these associated numbers may be the numbers that measure successively some changing physical quantity: temperature, pressure, velocity, and so forth. We represent this variable by some symbol x and say that under such and such conditions x has such and such values. As in a physical problem we may be concerned with more than one varying quantity, so in Analysis we may have to consider several variables x, y, z,... simultaneously.

We pass by a natural development of ideas from the 'unknowns' of Algebra through the 'current coordinates' of Analytical Geometry to the 'variables' of Analysis and appropriately choose our symbols for variables from the concluding letters of the alphabet.

In contradistinction to such variables, we have the numbers a, b, c,... that embody the data of the problem of which the variables are the subject. We distinguish these numbers as 'constants'. They are part of the framework within which the variables move. It may not be necessary to know which particular numbers a, b, c,... represent; they may even change their meaning as the inquiry proceeds. But as against the variables they are just constants. Often the distinction is sufficiently brought out by speaking of 'variables' as opposed to 'numbers'. We choose the symbols for constants, as for the 'knowns' of Algebra, from the beginning of the alphabet.

4. Continuous variation. Intervals

The completeness of the system of real number, whereby every point of a line can be represented arithmetically, allows the variable the possibility of *continuous variation*. By 'continuous variation' we mean, in geometrical language, that the moving point, in passing from P to Q on the scale of length, passes through every intermediate point. Analytically, x changes continuously from a to b through *every real number* between a and b. Physical quantities appear to us to undergo such 'continuous variation', at any rate in the more obvious phenomena of our experience. The wider implications of continuous variation are of much importance and are discussed at length in later chapters.

Continuous variation is not, of course, the only type of variation that is of interest to mathematicians. We may, for instance, restrict the values of the variable to the positive integers: such variables are characteristic of the Theory of Numbers. Equally so is continuous variation characteristic of the Calculus.

The aggregate of values assumed in given circumstances by a variable

is called the *domain* of the variable. The domain consisting of all real numbers between a and b is called the *interval* (a, b); such a domain is evidently specially appropriate to a variable that can undergo continuous variation.

If the end-points a, b themselves belong to the domain, i.e. if the domain is defined by the mixed inequality

$$a \leqslant x \leqslant b$$

(supposing $a < b$), then the interval is said to be *closed*. If the end-points are excluded from the domain, so that its definition is now

$$a < x < b,$$

then the interval is said to be *open*. The interval

$$a < x \leqslant b$$

is open at a and closed at b. I shall call such an interval *half-open* (or *half-closed*).

I shall write

(a, b) for the closed interval;

$)a, b($ for the open interval;

$)a, b)$ for the interval open at a and closed at b.

I shall also write (a, b) in speaking of an interval without special reference to its end-points.

By a 'neighbourhood' of a point is meant an *open* interval to which the point belongs. Thus $)a, b($ is a neighbourhood of c, if $a < c < b$. We say that a property holds 'in the neighbourhood' of a point, if it holds throughout *some* neighbourhood of the point.

Occasionally a property holds throughout a neighbourhood of a point A except actually at A itself. We then say that the property holds *near but not at* A; but that it holds *at and near* A, if A is not itself excluded. A neighbourhood of A with A excluded may be called an 'open' neighbourhood of A.

In extending these ideas to fields of more variables than one, we have a wide choice in our type of domain. Thus, in two dimensions, we may employ a domain with a curvilinear boundary, as, for instance, a circle. For our present purposes it is simplest to restrict ourselves to regions bounded by parallels to the coordinate-axes, i.e. to regions of the type

$$a_1 \leqslant x_1 \leqslant b_1, \quad a_2 \leqslant x_2 \leqslant b_2, \quad \ldots, \quad a_n \leqslant x_n \leqslant b_n.$$

In two dimensions this is, of course, merely a rectangle, but in higher space we soon exhaust our elementary vocabulary and I shall retain

the term 'interval' for such a region, denoting it by

$$(a_1, a_2,..., a_n;\ b_1, b_2,..., b_n).$$

Such an interval is closed. If all boundary-points are excluded, we have the fully open interval

$$a_1 < x_1 < b_1, \quad a_2 < x_2 < b_2, \quad ..., \quad a_n < x_n < b_n,$$

which I denote by $\quad)a_1, a_2,..., a_n;\ b_1, b_2,..., b_n($.

I do not indulge in further refinements.

The use of 'neighbourhood' can be carried over into fields of many variables.

5. Functional dependence

The operations of the Calculus turn fundamentally on the notion of a correspondence between two or more variables. In the simplest case we have only two variables x, y, and to each value of x corresponds a single value of y: to each value of y corresponds a single value of x. These conditions recall the geometrical one-to-one correspondence where, to ensure that the correspondence is homographic, we further stipulate that the correspondence be expressible by purely rational constructions.

In the analytical one-to-one correspondence we make no requirement of rationality and rely only on the unique tallying of values of x and y. We say in such case that the variables x, y are *functionally connected*: that y is a *one-valued function* of x, and similarly that x is a *one-valued function* of y.

In the field of real variables, for instance, the analytical relations

$$y = x^3 \quad \text{or} \quad y = \sinh x$$

define an analytical one-to-one correspondence (but not a geometrical one-to-one correspondence). On the other hand, the one-to-one correspondence

$$y = (ax+b)/(cx+d)$$

satisfies the geometrical as well as the analytical conditions.

If we waive the uniqueness of the tallying, so that to a single value of x may correspond one or more values of y, then y is a *many-valued function* of x: so x may be a many-valued function of y. The two variables are still 'functionally connected', and we denote the functional relationship analytically in any of the forms

$$f(x,y) = 0 \quad \text{or} \quad x = x(y) \quad \text{or} \quad y = y(x). \tag{1}$$

The domain of the values of x in the correspondence is called the *domain*

of definition of the function $y = y(x)$; similarly the domain of values of y is the domain of definition of the function $x = x(y)$.

In like fashion we may define a functional relation between many variables $x_1,..., x_n$ and may write it as $f(x_1,..., x_n) = 0$ or $x_1 = x_1(x_2,..., x_n)$, etc. The aggregate of the sets of values of $x_2,..., x_n$, i.e. of the 'points' $(x_2,..., x_n)$, constitutes the domain of definition of the function $x_1(x_2,..., x_n)$. If to each point $(x_2,..., x_n)$ corresponds a single value of x_1, then the function $x_1 = x_1(x_2,..., x_n)$ is single-valued.

This notion of functional dependence derives immediately from Physics. For there we are always endeavouring to establish a correspondence between the values of physical quantities as evidence of causal connexion and as material towards the formulation of the law of this connexion. We regard it as of the essence of this physical dependence between, say, n quantities that no one of them may assume values arbitrarily, when the remainder are fixed. For that would rather indicate physical *independence* and we should be convinced that our description of the phenomena could not be effected in terms of the n quantities alone.

Now, analytically, we have admitted many-valued functions and we ought not to exclude functions such as $y = \sin^{-1}x$ in which y may assume infinitely many values for each value of x. For the fundamental implication of functional dependence, that y is constant when x is constant, is still satisfied, even if the constants are infinitely many. But it would be repugnant to the idea of functional dependence if, when x were fixed, y could still enjoy continuous variation. We could then, for instance, claim $\sin(x_1+x_2)$ as a many-valued function of x_1 only, on the ground that, when x_1 is constant, the function also is constant, the constants being exactly the numbers of the interval $(-1, 1)$. I shall therefore exclude from the domain of definition of $y(x)$ a value of x to which corresponds a complete interval of values of y.

6. Functional definition

Nothing has yet been said as to how, in practice, we are to exhibit this correspondence of values that will define a function. Most simply we think of a table with corresponding values juxtaposed in adjacent columns.† Such, for instance, would be the tables of trigonometrical functions, if we could imagine them to be the only definition we possessed of these functions. But definition by a table can give only a finite

† I am thinking for the present of a functional relation between two variables x, y only.

set of values of either variable. Thus continuous variation is impossible and the processes of the Calculus are inapplicable.

The physicist may visualize a functional relation graphically in the form of a curve. This mental picture has its own value, but only as subordinate to an exact definition, which the curve itself cannot provide. For the curve, however drawn, lacks precision and yields information appropriate rather to Physics than to Mathematics.

The only adequate definition of a function for an infinity of values of either variable is by means of a rule: that is, of a process by which we can pass from a given value of x to the corresponding value of y or of a test by which we can determine what values of x, y can properly be paired. The process or the test ought, since we are concerned with Mathematics, to be stated in mathematical terms: it will depend presumably on the fundamental mathematical operations of addition, multiplication, and passage to the limit. We can give varied instances of functional definition:

y is the greatest integer in x; (2)

if x is a rational number, expressed in lowest terms as $(\pm p)/q$, $y = 1/q$: if x is irrational, $y = 0$; (3)

$$xy = x+y; \tag{4}$$

$$\sin(x+y) = x^2+y^2. \tag{5}$$

These examples illustrate certain characteristics of functional definition. Firstly, the domain of either variable need not be the whole of real number. In (4) we must exclude $x = 1$ and $y = 1$; in (5) the domains of x, y are certainly limited by the condition $x^2+y^2 \leqslant 1$. In (2), (3) the domain of x is the whole of real number, but the domains of y are respectively the integers, and the reciprocals of the positive integers together with zero.

Again, as in (3), the definition may change its form for different parts of the domain. It often rests between our convenience and our ingenuity how concise we make our definition. But it is of the essence of the functional notion that the value of y assigned to one value of x in no way influences the value of y to be assigned to any other value of x. We have an infinity of choice—so long as we can state it in finite terms.

7. Analytical expressions

In (2), (3) the functional definition is stated mainly in words; in (4), (5) it is in symbolic form: that is to say, the function is defined by an *analytical expression*. Historically, the functional idea grew up out of

such 'analytical expressions' formed of the elementary functions of analysis: x^n, $\log x$, $\sin x$, and their like; and at one time 'function' was held to be coextensive with 'analytical expression'. Such a point of view, moreover, ruled out a 'composite' function such as

$$y = \sin x \quad (x \geqslant 0) \qquad y = \tan x \quad (x \leqslant 0).$$

It would have been said that here were two functions, not one function.

But the development of the theory of Fourier series made this exclusive view untenable. For it was found that a Fourier series could provide an analytical expression, in the form of an infinite series of elementary functions, that represented in different parts of its domain distinct elementary functions.

Typically we can write the function (2), i.e. 'the greatest integer in x', in such a form. If we denote the function by one of the symbols usual for it, say $I(x)$, we have:

$$\left.\begin{aligned} &I(x) = x, \text{ if } x \text{ is zero or an integer;} \\ &I(x) = -\tfrac{1}{2}+x+\frac{\sin 2\pi x}{2\pi}+\frac{\sin 4\pi x}{4\pi}+\frac{\sin 6\pi x}{6\pi}+\ldots \text{ to } \infty, \\ &\qquad\qquad \text{if } x \text{ is not zero or an integer.} \end{aligned}\right\} \qquad (6)$$

The proof is not elementary and is postponed until chapter XIV.†

More simply we may write

$$I(x) = x-\tfrac{1}{2}+\frac{1}{\pi}\tan^{-1}(\cot x),$$

where $\tan^{-1}$ denotes the 'principal value' of the inverse tangent, i.e. that value lying in the half-open interval $)-\tfrac{1}{2}\pi, \tfrac{1}{2}\pi)$, or again

$$I(x) = x-\tfrac{1}{2}-\frac{2}{\pi}\tan^{-1}\{\tan(\tfrac{1}{2}\pi x \pm \tfrac{1}{4}\pi)\},$$

where the upper or the lower sign is to be taken in $\pm$ according as $I(x)$ is odd or even. With this convention, $\tan(\tfrac{1}{2}\pi x \pm \tfrac{1}{4}\pi)$ does not numerically exceed unity, and we may therefore expand the inverse tangent in powers of its argument, getting

$$\begin{aligned} I(x) = x-\tfrac{1}{2}-\frac{2}{\pi}\tan(\tfrac{1}{2}\pi x \pm \tfrac{1}{4}\pi)+\frac{2}{3\pi}\tan^3(\tfrac{1}{2}\pi x \pm \tfrac{1}{4}\pi)-& \\ -\frac{2}{5\pi}\tan^5(\tfrac{1}{2}\pi x \pm \tfrac{1}{4}\pi)+\ldots \text{ to } \infty,& \qquad (7) \end{aligned}$$

where the upper or the lower signs are to be taken according as $I(x)$ is odd or even.

We are thus driven to fix attention on the correspondence of values

† Chapter XIV, § 4 (20).

as the essential fact of functionality and to regard the analytical expression, if one can be found, not as the function itself but merely as a convenient representation of the function. We take this view the more readily, since, in practice, we frequently require for the function, not just any analytical expression, but an expression of some particular type: for instance, an expression in power-series. Consider in this way the function (4) above. Solving for y, we may write

$$y(x) = (1-x^{-1})^{-1} = 1+x^{-1}+x^{-2}+\ldots\text{to }\infty.$$

This power-series converges only in the region $|x| > 1$. But we may also write

$$y(x) = -x(1-x)^{-1} = -x-x^2-x^3-\ldots\text{to }\infty,$$

which is valid in the region $|x| < 1$. The points $x = \pm 1$ are excluded from both these regions of convergence: the point $x = 1$ reasonably so, since $y(x)$ is undefined there. We can cover the point $x = -1$, if, for instance, we write

$$y(x) = -\frac{x}{2}\left(1-\frac{x+1}{2}\right)^{-1} = -\frac{x}{2}\left\{1+\frac{x+1}{2}+\left(\frac{x+1}{2}\right)^2+\ldots\text{to }\infty\right\},$$

where $|x+1| < 2$.

The general question of representation of functions by power-series is discussed at length in chapters VIII and XIV. But enough has now been said to show, on the one hand, that we cannot generally expect a single power-series to provide a complete representation of a function; and, on the other hand, that power-series giving a representation over a restricted domain exist in endless variety. We fall back on the view that power-series and analytical expressions generally are capable of illustrating aspects of the function (just as a sketch or a diagram will illustrate some aspect of a solid object) but that the essential function is the correspondence of values, however arrived at.

8. Implicit and explicit functions

There is one other point to make. In (2), (3) the function is defined by a *rule* that tells us how to proceed from a given value of x to the corresponding value of y. We say that $y(x)$ is an 'explicit' function, that x is the 'independent' variable, and that y is the 'dependent' variable. In (4), (5) we are given a *test* to determine whether two specified numbers can stand together as respective values of x and y in the functional correspondence. We say that the functional relations (4), (5) are 'implicit'. We can at once write (4) in the explicit form

$$y = x/(x-1),$$

in which x is the independent variable and y the dependent variable; or in the explicit form

$$x = y/(y-1),$$

in which y is the independent variable and x the dependent variable. But in (5) it is a matter of much difficulty, given a value of x, to determine the corresponding value or values of y: and so generally with implicit functions.

The distinction between implicit and explicit functions is not inherent in the function itself; it is concerned with our convenience and depends merely on the mode of defining the function. In the function defined by a table, $y(x)$ is explicit, if the table is arranged 'alphabetically' in x, i.e. so that we enter the table for values of x and read from it values of y; $x(y)$ is explicit, if we enter the table for values of y, reading from it values of x. But the function is defined implicitly, if the table is arranged higgledy-piggledy in x and y alike.

In an explicit function the independent variable is often called the 'argument' of the function. If x, y are functionally related, we may call the two functions $y(x)$, $x(y)$ 'inverse' functions and we may write

$$y = f(x), \qquad x = f^{-1}(y).$$

If x, y are connected by an implicit relation $f(x,y) = 0$, it may not be easy or possible to define either of the explicit functions $y(x)$ or $x(y)$, but it may be possible to express x, y as explicit functions of a third variable t in forms

$$x = x(t), \qquad y = y(t)$$

satisfying the identity $$f\{x(t), y(t)\} = 0.$$

We call this third variable t a 'parameter', saying that the function $f(x,y) = 0$ is defined 'parametrically' by the equations $x = x(t)$, $y = y(t)$. The notion is geometric in origin, (x,y) being, characteristically, a moving point that traces out the curve $f(x,y) = 0$ and t the time of travel.

In the foregoing discussion we have been thinking of functional relations between two variables. But it is not difficult to see that the ideas and the terminology can be extended to functions of many variables.

9. Sequences of variable terms

If, in a function of m arguments $f(x_1,..., x_m)$, one argument x_1 can have the positive integers for its domain, we can thereby define the *sequence*

$$\{f(n, x_2,..., x_m)\}.$$

If $m > 1$, this is a sequence of variable terms. If the sequence con-

verges at a set of points $(x_2,..., x_m)$, the values of the corresponding limits define a new function of these points

$$f(x_2,..., x_m),$$

whose domain of definition is the region of convergence. We call this new function the 'limit-function' of the sequence. Most frequently in practice the sequence is given in the form of an infinite series. The 'limit-function' of the sequence then becomes the 'sum-function' of the series.

One of the most fertile methods of introducing new and interesting functions into Analysis is as the limit-functions or the sum-functions of sequences or series of elementary functions, and it is worth while to give some attention to functions so defined. I confine myself to the simplest case of a single continuous variable and consider a function $f(x)$ defined as the limit of a sequence $s(n,x)$.

Let $x = \xi$ be a point of convergence of the sequence. By this we mean, of course, that, for any positive ϵ, the inequality

$$|s(n,\xi)-s(\xi)| < \epsilon \tag{8}$$

can be secured by some inequality

$$n > N. \tag{9}$$

Now N depends on ϵ, in general increasing as ϵ diminishes; it depends also on ξ. We may put this in evidence by rewriting (9) as

$$n > N(\epsilon,\xi). \tag{10}$$

There is a least N that makes (10) only just sufficient to secure (8). By choosing this least value we make $N(\epsilon,\xi)$ precise and so define the function

$$N(\epsilon,x)$$

over positive values of ϵ and values of x in the domain of convergence of $s(n,x)$.

Now in practice analytical operations on $s(x)$ must usually be effected by means of similar operations on $s(n,x)$. It is therefore important to know how closely $s(n,x)$ represents $s(x)$, i.e. how rapidly the sequence converges, in different parts of its domain.

The function $N(\epsilon,x)$, for fixed ϵ and varying x, gives us a measure of the rapidity, or rather of the slowness of convergence. If, for every ϵ, this function of x is bounded throughout some region,† i.e. if

$$N(\epsilon,x) < N_0(\epsilon)$$

† $N(\epsilon, x)$ as a function of ϵ is not, in general, bounded, since N diverges as $\epsilon \to 0$.

throughout this region, then $N_0(\epsilon)$ may replace $N(\epsilon, \xi)$ in (10) and so

$$|s(n,\xi)-s(\xi)| < \epsilon$$

is secured by

$$n > N_0(\epsilon)$$

for every ξ in this region.

10. Uniform convergence

In this case $s(n,x)$ represents $s(x)$ to within ϵ throughout this region and so can replace it (we may expect) in any analytical operations. The convergence is then said to be 'uniform over the region': the region is a 'region of uniform convergence'.

In the contrary case, for sufficiently small standards of accuracy ϵ, we can find no $s(n,x)$ capable of representing $s(x)$ with this accuracy throughout the region. The convergence is 'non-uniform' in the region.

It must never be forgotten that the presence of the variable is essential to the notion of uniform convergence and that the word 'uniform' is meaningless except with reference to some domain of the variable, expressed or implied.

The condition of 'uniform convergence to $s(x)$' is, as we have just seen, that, over some domain of x,

$$|s(n,x)-s(x)| < \epsilon, \quad \text{if} \quad n > N(\epsilon). \tag{11}$$

If we have no independent knowledge of $s(x)$, it is more convenient to use the 'general condition of uniform convergence' in which $s(x)$ does not appear. It is that, over some domain of x and for every positive integer p,

$$|s(n,x)-s(n+p,x)| < \epsilon, \quad \text{if} \quad n > N(\epsilon). \tag{12}$$

Clearly (11) implies (12), for by (11), if $n > N(\frac{1}{2}\epsilon)$,

$$|s(n,x)-s(x)| < \tfrac{1}{2}\epsilon \quad \text{and} \quad |s(n+p,x)-s(x)| < \tfrac{1}{2}\epsilon$$

for every positive integer p. Hence

$$\begin{aligned}|s(n,x)-s(n+p,x)| &\leqslant |s(n,x)-s(x)|+|s(n+p,x)-s(x)| \\ &< \epsilon.\end{aligned}$$

Conversely (12) implies (11), for (12) shows that $s(n,x)$ converges for each x in the domain and therefore the limit-function $s(x)$ is defined throughout the domain. In (12) we may then take the limit $p \to \infty$, which gives

$$|s(n,x)-s(x)| \leqslant \epsilon, \quad \text{if} \quad n > N(\epsilon),$$

and so, to secure a pure inequality,

$$|s(n,x)-s(x)| < \epsilon, \quad \text{if} \quad n > N(\tfrac{1}{2}\epsilon).$$

The equivalence between (11), (12) is thus exact.

We can restate these two conditions in a form applicable to series.

If $S_n(x)$ is the sum to n terms of a series that converges to the sum-function $S(x)$, then this convergence is uniform over a domain of x throughout which

$$|S_n(x)-S(x)| < \epsilon, \quad \text{if} \quad n > N(\epsilon), \tag{13}$$

or, equivalently,

$$|S_n(x)-S_{n+p}(x)| < \epsilon, \quad \text{if} \quad n > N(\epsilon), \tag{14}$$

for every positive integer p.

11. The uniform convergence of power-series

The principles of uniform convergence apply with especial simplicity to power-series, in virtue of the following theorem:

(15) *A power-series $\sum a_n x^n$, if it converges at $x = \xi$, converges uniformly over the closed interval* $(0, \xi)$.

Write $x \equiv \xi v$, $S_n(x) \equiv \sum_1^n a_n x^n$, $s_n \equiv S_n(\xi)$. Then

$$\begin{aligned} &S_{n+p}(x)-S_n(x) \\ &= a_{n+1}x^{n+1}+a_{n+2}x^{n+2}+\ldots+a_{n+p}x^{n+p} \\ &= (s_{n+1}-s_n)v^{n+1}+(s_{n+2}-s_{n+1})v^{n+2}+\ldots+(s_{n+p}-s_{n+p-1})v^{n+p} \\ &= (s_{n+1}-s_n)(v^{n+1}-v^{n+2})+(s_{n+2}-s_n)(v^{n+2}-v^{n+3})+\ldots+(s_{n+p}-s_n)v^{n+p}. \end{aligned}$$

Now $0 \leqslant v \leqslant 1$, if x lies in the closed interval $(0, \xi)$, and so $v^{n+r}-v^{n+r+1}$ $(r = 1,\ldots, p-1)$ is not negative. Hence, since the modulus of a sum does not exceed the sum of the moduli,

$$\begin{aligned} |S_{n+p}(x)-S_n(x)| \leqslant\ &|s_{n+1}-s_n|(v^{n+1}-v^{n+2})+ \\ &+|s_{n+2}-s_n|(v^{n+2}-v^{n+3})+\ldots+|s_{n+p}-s_n|v^{n+p}. \end{aligned}$$

Since the series s_n converges, we can find $N(\epsilon)$ such that $n > N(\epsilon)$ implies $|s_{n+r}-s_n| < \epsilon$ for every r. Thus

$$\begin{aligned} |S_{n+p}(x)-S_n(x)| &< \epsilon(v^{n+1}-v^{n+2}+v^{n+2}-v^{n+3}+\ldots+v^{n+p}) \\ &= \epsilon v^{n+1} \\ &\leqslant \epsilon, \quad \text{if} \quad n > N(\epsilon). \end{aligned}$$

Now $N(\epsilon)$ is independent of x, since it has been determined solely from a knowledge of the convergent series of constants $\sum a_n \xi^n$. The convergence is therefore uniform over the closed interval $(0, \xi)$.

As regards the ordinary convergence of the power-series $\sum a_n x^n$ at a point $x = \xi$, we know† that the series is convergent or non-convergent according as

$$\overline{\lim}\, |a_n|^{1/n}|\xi| < 1 \quad \text{or} \quad > 1,$$

and that, if this limit is equal to unity, either alternative is possible.

† Cf. Bromwich, *Theory of Infinite Series* (1908), §§ 10, 50.

If we write $$\rho \equiv \underline{\lim}\, |a_n|^{-1/n},$$
the region of convergence of the power-series is the interval $(-\rho, \rho)$ which may be open, closed, or half-open. Thus the several series
$$\sum x^n, \qquad \sum x^n/n, \qquad \sum x^n/n^2$$
have respectively the intervals of convergence
$$)-1, 1(, \qquad (-1, 1(, \qquad (-1, 1).$$
For $\sum n!\,x^n$, $\rho = 0$ and the interval of convergence reduces to the isolated point $x = 0$; such a series is of little interest at this stage. On the other hand $\sum x^n/n!$ converges for all real values of x.†

Now if a sequence converges uniformly over each of a finite set of domains $D_1, \ldots, D_n$, it converges uniformly over the aggregate of these domains. For, if $N(\epsilon, x)$, regarded as a function of x only, is bounded in each domain, it is bounded in the aggregate of the domains.

Thus, if the power-series converge for $x = \pm\rho$, it converges uniformly over each of the closed intervals $(0, \rho)$, $(0, -\rho)$ and therefore over the closed interval $(-\rho, \rho)$. Hence, if the interval of ordinary convergence is closed, it is also the interval of uniform convergence.

Suppose now that the series converge at $-\rho$ but not at ρ. It still converges at any ρ_1, where $0 < \rho_1 < \rho$, and therefore it converges uniformly over the closed interval $(-\rho, \rho_1)$. Similarly, if the interval of ordinary convergence is the open interval $)-\rho, \rho($, there is uniform convergence over any closed interval $(-\rho_2, \rho_1)$, where $0 < \rho_1, \rho_2 < \rho$. We may sum up these results concisely by saying that:

(16) *A power-series converges uniformly over any closed interval that belongs to the interval of convergence.*

We cannot state a more precise region of uniform convergence, since, as we shall see later, the interval of uniform convergence of a power-series can always be closed and therefore cannot coincide with an open or half-open interval of ordinary convergence. There is thus no *complete* interval of uniform convergence except when the interval of ordinary convergence is closed.

12. Non-uniform convergence

When the condition of uniform convergence is not satisfied in respect of some region, we may say that the convergence is 'non-uniform' in that region. It is convenient to determine an explicit analytical condition for non-uniform convergence.

† If we like, we may say that it converges throughout the interval $(-\infty, \infty)$, essentially open.

The condition for uniform convergence over a domain is, in precise terms, that

> for *every* ϵ we can find *some* $N(\epsilon)$ such that $|s(n,x)-s(n+p,x)| < \epsilon$ for *every* x in the domain, for *every* p, and for *every* n greater than $N(\epsilon)$.

The condition of uniform convergence to $s(x)$ is similarly that

> for *every* ϵ we can find *some* $N(\epsilon)$ such that $|s(n,x)-s(x)| < \epsilon$ for *every* x in the domain and for *every* n greater than $N(\epsilon)$.

To contradict these conditions we interchange the words 'some' and 'every' and reverse the central inequality. Thus the condition for non-uniform convergence in a region is that

> for *some* positive ϵ and *every* N, there is *some* $n(N)$ greater than N with which can be associated *some* $x(n)$ of the region and *some* positive integer $p(n)$ to secure the inequality
>
> $$|s(n,x)-s(n+p,x)| > \epsilon.$$

In other words,

(17) *We can specify an ϵ such that, by suitable choice of $x(n)$ in the region and of $p(n)$, both depending on n, the inequality*

$$|s(n,x)-s(n+p,x)| \geqslant \epsilon$$

is satisfied for arbitrarily large values of n.

It is to be remembered that this condition admits the possibility of non-convergence as well as of non-uniform convergence in the region.

In like fashion, if there is non-uniform convergence to $s(x)$,

(18) *We can specify an ϵ such that, by suitable choice of $x(n)$ in the region, the inequality*

$$|s(n,x)-s(x)| \geqslant \epsilon$$

is satisfied for arbitrarily large values of n.

We should remember that (18) admits the possibility not only of non-convergence in the region but also of convergence to a limit-function other than $s(x)$.

As an example of non-uniform convergence consider the power-series

$$S(x) \equiv \sum x^n$$

in the open interval $)-1, 1($. (19)

Take $\epsilon < \epsilon_1 < 1$, $p = 1$, $x = \pm \sqrt[n+1]{\epsilon_1}$, so that x lies in $)-1, 1($. Then

$$|S_{n+1}(x)-S_n(x)| = |x|^{n+1} = \epsilon_1 > \epsilon,$$

the inequality holding for all values of n.

The convergence (which we know to extend over the open interval) is therefore, by (17), non-uniform over that interval. This result throws

light on what has been said regarding the regions of uniform convergence of a power-series.

As a second example consider, near $x = 0$,

$$s(n,x) \equiv \frac{nx}{1+n^2x^2}. \tag{20}$$

Here $s(x) = 0$ for every x. In particular, writing $x = \pm 1/n$, we have

$$|s(n,x)-s(x)| = \tfrac{1}{2},$$

so that with $\epsilon = \frac{1}{2}$ we see from (18) that the convergence is non-uniform over any region containing points $x = \pm 1/n$ i.e. over any neighbourhood of the origin.

We should observe that the non-uniformity of convergence in the neighbourhood of this point is not accounted for by a failure of convergence *at* the point: the sequence in fact converges to zero at the point. We may call such a point rather loosely 'a point of non-uniform convergence'.

On the other hand, we saw that the series $\sum x^n$ did not converge uniformly over any interval $(1-\epsilon, 1($. In this case the 'point of non-uniform convergence' $x = 1$ is also a point of non-convergence.

We can show further by an example that

(21) *A point of non-convergence is not necessarily a point of non-uniform convergence.*

For write
$$s(n,x) \equiv \{I(x)\}^n, \tag{22}$$
where $I(x)$ denotes as before the greatest integer in x.

Then $I(2) = 2$, but $I(x) = 1$ in the interval $(1, 2($. Thus $s(n,x)$ diverges at $x = 2$ and converges in $(1, 2($; in this half-open interval we have, in fact, $s(n,x) = 1 = s(x)$. The convergence is therefore uniform over the half-open interval $(1, 2($, despite the fact that there is divergence at $x = 2$.

We must not, however, form the idea that convergence must always be uniform except in the neighbourhood of isolated points.

As an example of a sequence that *converges everywhere but converges uniformly over no interval* write

$$\begin{aligned} s(n,x) &\equiv \sin(n!x\pi) \quad (x \text{ rational}),\\ s(n,x) &\equiv 0 \qquad\qquad\quad (x \text{ irrational}). \end{aligned} \tag{23}$$

If x, when rational, is expressed in lowest terms as p/q, then $n!x$ is integral when $n \geqslant q$, and so $s(n,x)$ vanishes. Thus the limit-function exists everywhere and is everywhere zero.

Consider a rational point $x = p/q$, where q is prime. Then, by Wilson's theorem,† $(q-1)!+1$ is a multiple of q, and so, if $n = q-1$,

$$|s(n,x)-s(x)| = |s(n,x)| = |\sin x\pi|.$$

If x is restricted to an interval (a,b), where $0 < a < b < \frac{1}{2}$, we have

$$|\sin x\pi| > \sin a\pi.$$

Now, so long as $q > (b-a)^{-1}$, we can find a point p/q in the interval (a,b): and we can find primes q arbitrarily great. Hence, taking

$$\epsilon = \sin a\pi, \qquad n = q-1, \qquad x = p/q, \qquad q \text{ prime},$$

we satisfy the condition (18) of non-uniform convergence.

We have thus proved that there is uniform convergence over no part of the interval $(0, \frac{1}{2})$; and we can deal similarly with any other finite interval.

13. Trigonometric series

We consider here certain trigonometric series of the type

$$\sum a_n \sin nx \quad \text{or} \quad \sum a_n \cos nx.$$

They have the form of Fourier's series but are not properly to be called Fourier's series unless their coefficients have certain intimate relations with a certain sum-function, real or supposed. To limit the discussion I restrict the coefficients a_n to be non-negative.

Now it is known from Abel's lemma‡ that, in the particular case in which the sequence $\{a_n\}$ converges *monotonically* to zero, the partial remainder $S_n(x)-S_{n+p}(x)$ in either series is numerically less than

$$|a_n \operatorname{cosec} \tfrac{1}{2}x|.$$

There is thus uniform convergence over any interval that excludes the points $x = 2k\pi$, where k is integral or zero. The series are periodic with period 2π. It is thus sufficiently general to consider convergence in the neighbourhood of $x = 0$.

The substitution $x = 0$ reduces $\sum a_n \sin nx$ to a sum of zeros; it therefore converges to zero. Thus the sine-series converges everywhere.

The substitution $x = 0$ reduces $\sum a_n \cos nx$ to the series of positive constants, $\sum a_n$. If this latter series converge, then

$$|a_{n+1}\cos(n+1)x+\dots+a_{n+p}\cos(n+p)x| < a_{n+1}+\dots+a_{n+p}$$
$$< \epsilon, \quad \text{if } n > \text{some } N(\epsilon),$$

and so, in this case, the cosine-series converges everywhere and converges uniformly over every interval.

† Cf. Chrystal, *Algebra*, 2 (1900), 553. ‡ Cf. Bromwich, loc. cit. § 20, Ex.

Alternatively, if $\sum a_n$ diverge, then $\sum a_n \cos nx$ diverges at $x = 0$. Moreover, since $\sum a_n$ diverges, we can find ϵ and $p(n)$ such that

$$|a_{n+1}+...+a_{n+p}| > \epsilon$$

for arbitrarily large values of n. If further we take $x = \pi/3(n+p)$, we have

$$|a_{n+1}\cos(n+1)x+...+a_{n+p}\cos(n+p)x| > \tfrac{1}{2}|a_{n+1}+...+a_{n+p}| > \tfrac{1}{2}\epsilon$$

for arbitrarily large values of n. Hence by (17) there is non-uniform convergence in the neighbourhood of $x = 0$. This disposes of the cosine-series for the particular case we are considering, namely, that of a_n converging to zero monotonically.

The discussion of the uniform convergence of the sine-series under similar conditions is less simple. It turns on the following pair of inequalities:

(24) *If a_n converge monotonically to zero, then*

(i) *for every p and every x*

$$|a_{n+1}\sin(n+1)x+...+a_{n+p}\sin(n+p)x| < (\pi+1)A_n,$$

where A_n is the greatest term (or the upper bound) of the infinite set

$$na_n,\ (n+1)a_{n+1},...;$$

(ii) *for every p and some $x(p)$*

$$|a_{n+1}\sin(n+1)x+...+a_{n+p}\sin(n+p)x| > \frac{p(n+1)}{n+p}a_{n+p}.$$

For (i), it is sufficient to take $0 < x < \pi$. Then†

$$\operatorname{cosec} \tfrac{1}{2}x < \pi/x.$$

Let r be the greatest integer in π/x, and suppose at first that $n < r < n+p$. Then

$$\begin{aligned}|a_{n+1}\sin(n+1)x+...+a_r\sin rx| &< \{(n+1)a_{n+1}+...+ra_r\}x\\ &< (r-n)\pi A_n/r\\ &< \pi A_n.\end{aligned}$$

But by Abel's lemma

$$\begin{aligned}|a_{r+1}\sin(r+1)x+...+a_{n+p}\sin(n+p)x| &< a_{r+1}\operatorname{cosec}\tfrac{1}{2}x\\ &< \pi a_{r+1}/x\\ &< (r+1)a_{r+1} \leqslant A_n.\end{aligned}$$

Hence (i) is proved, if $n < r < n+p$; and it is evidently true *a fortiori*, if either $r \leqslant n$ or $r \geqslant n+p$.

For (ii), we have at once that the set

$$\sin(n+1)x,\ \sin(n+2)x,...,\ \sin(n+p)x$$

† For a proof of this inequality see, for instance, chapter IV (43).

is ascending monotonic, if $(n+p)x \leqslant \frac{1}{2}\pi$, and by hypothesis $\{a_n\}$ is descending monotonic. Hence

$$|a_{n+1}\sin(n+1)x+\ldots+a_{n+p}\sin(n+p)x| > pa_{n+p}\sin(n+1)x$$
$$> \frac{p(n+1)}{n+p}a_{n+p},$$

if we take the value of $x(p)$ to be that given by

$$(n+p)x = \tfrac{1}{2}\pi.$$

From the inequalities (24) we deduce† that

(25) *If a_n converge monotonically to zero, the condition $na_n \to 0$ is both necessary and sufficient for the uniform convergence over every interval of the sine-series*

$$\sum a_n \sin nx.$$

The condition $na_n \to 0$ is sufficient, for then $A_n < \epsilon/(\pi+1)$, if $n >$ some $N(\epsilon)$, and so, by (24) (i),

$$|a_{n+1}\sin(n+1)x+\ldots+a_{n+p}\sin(n+p)x| < \epsilon$$

for every p, every x, and every $n > N(\epsilon)$, which is the condition of uniform convergence over every interval.

Conversely, if there is uniform convergence over every interval, then, by (24) (ii), we must have for every p

$$\frac{p(n+1)}{n+p}a_{n+p} < \epsilon, \quad \text{if } n > \text{some } N(\epsilon).$$

Write successively $p = n$, $n+1$ and we have

$$2na_{2n} < 4\epsilon,\ (2n+1)a_{2n+1} < 4\epsilon, \quad \text{if } n > N(\epsilon),$$

which ensures that $na_n \to 0$.

WORKED EXAMPLE

Prove that the aggregate of real numbers cannot be put into one-to-one correspondence with the aggregate of positive integers; nor the aggregate of real functions into one-to-one correspondence with the aggregate of real numbers. (CANTOR.)‡

To put the real numbers into one-to-one correspondence with the positive integers is, in other words, to arrange them in a sequence. But we can show that it is impossible to arrange even the real numbers of the interval $(0, 1)$ completely in a sequence.

Let us write the numbers of this interval as decimal fractions. Those rational numbers of the interval whose denominators contain no prime factors other than 2, 5 can be written either as terminated decimals, that is to say, as recurring decimals with recurring part $\dot{0}$, or alternatively as recurring decimals with recurring part $\dot{9}$. If we agree to exclude $\dot{9}$ as a permissible recurring part, then every number, rational or irrational, of the interval $(0, 1)$ can be written uniquely as

† Chaundy and Jolliffe, *Proc. London Math. Soc.* (2) 15, (1916), 214–16.

‡ Cf. *Gesammelte Abhandlungen*, § 8, 278–81.

an infinite decimal. Suppose then that these numbers of this interval can be arranged completely in the sequence

$$u_1 = 0{\cdot}a_{11}a_{12}a_{13}\cdots$$
$$u_2 = 0{\cdot}a_{21}a_{22}a_{23}\cdots$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$
$$u_n = 0{\cdot}a_{n1}a_{n2}a_{n3}\cdots$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

Consider now the number

$$u' = 0{\cdot}b_1 b_2 b_3 \ldots$$

where, for every n,

$$b_n = 9-a_{nn}.$$

Then this number u', which clearly belongs to the interval $(0, 1)$, finds no place in the given sequence. For suppose it comes nth, say, in the sequence, i.e. is equivalent to u_n. Then, since the numbers of the interval have, under our convention, been written uniquely as infinite decimals, u_n and u' must be identical in *every* digit. But their nth digits are respectively

$$a_{nn} \quad \text{and} \quad b_n,$$

and these cannot be the same, since, by definition, $b_n = 9-a_{nn}$. Thus u' has been omitted from the given sequence, and, by a similar argument, no sequence can be found to contain all numbers of the interval $(0, 1)$.

We may prove in a similar way that the real functions cannot completely be brought into one-to-one correspondence with the real numbers. It is clearly sufficient to prove this for a restricted class of such functions, namely, those defined for *all* real values of their arguments.

Suppose then that the correspondence has been established. We can write $f_a(x)$ for the real function that corresponds to the real number a. Now consider the real function $F(x)$ defined, for every real x, as

$$F(x) = \{f_x(x)\}^2+1.$$

This function has been omitted from the correspondence. For, if it correspond to the real number a, i.e. if it be the real function $f_a(x)$, then we must have for every real x,

$$f_a(x) = \{f_x(x)\}^2+1,$$

and in particular, for $x = a$,

$$f_a(a) = \{f_a(a)\}^2+1.$$

But this is impossible, since $f_a(x)$ is real for every real x. Thus the function $F(x)$ has been omitted from the correspondence. In other words, the correspondence is incomplete and so, by a similar argument, will be every other such one-to-one correspondence that can be devised.

EXAMPLES I

1. For what domains of values of x are the following defined as real functions:

$$\sin^{-1}(\sinh x), \qquad \cos^{-1}(\cosh x), \qquad \tan^{-1}(\tanh x),$$
$$\sinh^{-1}(\sin x), \qquad \cosh^{-1}(\cos x), \qquad \tanh^{-1}(\tan x),$$
$$\cosh^{-1}(\tan x), \qquad \cos^{-1}(\tanh x)\,?$$

2. Sketch the graphs of the sum-functions of the infinite series:

(i) $\sin x+\frac{1}{3}\sin 3x+\frac{1}{5}\sin 5x+\dots$,

(ii) $x+\sin 2x+\frac{1}{2}\sin 4x+\frac{1}{3}\sin 6x+\dots$,

(iii) $\dfrac{\sin x}{x}+\dfrac{\sin 2x}{2x}+\dfrac{\sin 3x}{3x}+\dots$,

(iv) $\cos x+\dfrac{1}{2}\dfrac{\cos^3 x}{3}+\dfrac{1.3}{2.4}\dfrac{\cos^5 x}{5}+\dots$,

(v) $x+\cos x+\dfrac{1}{2}\dfrac{\cos^3 x}{3}+\dfrac{1.3}{2.4}\dfrac{\cos^5 x}{5}+\dots$,

(vi) $\sin x+\dfrac{1}{2}\dfrac{\sin^3 x}{3}+\dfrac{1.3}{2.4}\dfrac{\sin^5 x}{5}+\dots$,

(vii) $(1-\frac{1}{2}\pi)\sin x+\dfrac{1}{2}\dfrac{\sin^3 x}{3}+\dfrac{1.3}{2.4}\dfrac{\sin^5 x}{5}+\dots$,

(viii) $(\cos x-\sin x)+\left(\dfrac{1}{2}\dfrac{\cos^3 x}{3}-\dfrac{\sin 2x}{2}\right)+\left(\dfrac{1.3}{2.4}\dfrac{\cos^5 x}{5}-\dfrac{\sin 3x}{3}\right)+\dots$.

3. If $s(n,x)$, $\sigma(n,x)$ converge uniformly over a certain interval, show that their sum $s(n,x)+\sigma(n,x)$ converges uniformly over the interval; and that their product $s(n,x).\sigma(n,x)$ converges uniformly over the interval, if $s(n,x)$, $\sigma(n,x)$ are bounded in the interval.

4. If
$$s(n,x)=x(1+1/n)$$
and
$$\sigma(n,x)=\begin{cases} q+1/n & (x \text{ rational } = p/q),\\ 1/n & (x \text{ irrational}),\end{cases}$$
show that $s(n,x)$, $\sigma(n,x)$ each converge uniformly over any finite interval, but that their product $s(n,x).\sigma(n,x)$ converges uniformly over no interval.

5. Establish the comparison test for uniform convergence, i.e. that $s(n,x)$ converges uniformly over any interval in which
$$|s(n,x)-s(n+1,x)| < |\sigma(n)-\sigma(n+1)|,$$
where $\sigma(n)$ is a convergent monotonic sequence of constants.

By considering
$$s(n,x)=x^n+(-1)^n,$$
$$\sigma(n)=\xi^n/(1-\xi^2) \qquad (0<\xi<1)$$
over the interval $|x|<\xi$, or otherwise, show that a 'comparison test'
$$|s(n,x)-s(n+2,x)| < |\sigma(n)-\sigma(n+2)|$$
is insufficient even for ordinary convergence.

6. Prove that, as $n\to\infty$, $(1+x/n)^n$ converges to e^x uniformly over any finite interval, and $n(\sqrt[n]{x}-1)$ converges to $\log x$ uniformly over any interval $0<a<x<b$.

7. Determine the intervals of ordinary convergence, absolute convergence, and uniform convergence of the following power-series:

(i) $x-\frac{1}{2}x^2+\frac{1}{3}x^3-\frac{1}{4}x^4+\dots$,

(ii) $x^2-\frac{1}{2}x^4+\frac{1}{3}x^6-\frac{1}{4}x^8+\dots$,

(iii) $x+\frac{1}{2}x^2-\frac{2}{3}x^3+\frac{1}{4}x^4+\frac{1}{5}x^5-\frac{2}{6}x^6+\dots$,

(iv) $x+(1+\frac{1}{2})\dfrac{x^2}{2}+(1+\frac{1}{2}+\frac{1}{3})\dfrac{x^3}{3}+\dots$.

8. Examine for convergence and uniform convergence the sequences

$$\text{(i)}\ \frac{nx}{1+nx^2+n^2x^2},\qquad \text{(ii)}\ \frac{n^2x^3}{1+nx^2+n^3x^2},\qquad \text{(iii)}\ \frac{nx}{1+nx^2+n^2x^3},$$

$$\text{(iv)}\ \frac{x^a n^{1-b}}{1+nx^2}\quad (a,b>0),\qquad \text{(v)}\ x^a n^b(1-x)^n\quad (a,b>0,\ 1\geqslant x>0).$$

9. Examine for convergence and uniform convergence the infinite series

$$\sum (x^2+n^2)^{-1},\qquad \sum \{(x^2+n)^{-1}-n^{-1}\},$$
$$\sum (x^2+n^2)^{-p}\qquad (0<p<1)$$

10. If $a_n \to 0$ monotonically, show that the series

$$\sum a_n \sin \tfrac{1}{2}x \sin nx$$

converges uniformly over any finite interval.

11. Examine the following series for convergence and uniform convergence:

$$\text{(i)}\ \sum \frac{1}{n}\cos nx,\qquad \text{(iii)}\ \sum \frac{1}{n}\sin n\alpha \sin nx,$$

$$\text{(ii)}\ \sum \frac{1}{n}\cos n\alpha \sin nx,\qquad \text{(iv)}\ \sum \frac{1}{n}\cos^2 n\alpha \sin nx,$$

$$\text{(v)}\ \sum \frac{1}{n}\sin n\alpha \sin n\beta \sin nx.$$

12. Discuss the convergence and uniform convergence of

$$\sum \frac{\sin^2 nx}{n},\qquad \sum \frac{(-)^n \sin^2 nx}{n},\qquad \sum \frac{|\sin nx|}{n},$$

and show that

$$\sum \frac{\sin nx \cos ny}{n}$$

converges everywhere but is not uniformly convergent over any region cut by any of the lines

$$x = 2m\pi \pm y,$$

where m is a positive integer or zero.

13. Discuss the convergence and uniform convergence of the series

$$\sum_0^\infty \cos^n x \cos nx,\qquad \sum_0^\infty \cos^n x \sin nx,$$

and sketch the graph of the corresponding sum functions.

14. Prove that, when π/x is an *odd* integer, the infinite series

$$\tan x + \tfrac{1}{2}\tan 2x + \tfrac{1}{3}\tan 3x + \ldots \quad\text{and}\quad \sin x + \tfrac{1}{2}\sin 2x + \tfrac{1}{3}\sin 3x + \ldots$$

converge to equal sums.

15. If
$$S(x) = \frac{\sin \pi x}{\pi} + \frac{\sin 2\pi x}{2\pi} + \frac{\sin 3\pi x}{3\pi} + \ldots \text{ to infinity},$$

prove that, for every x,

$$S(x) = S(2x)+S(1-x).$$

16. Prove that
$$\sin x + \tfrac{1}{2}\sin 2x + \ldots + \frac{1}{n}\sin nx$$

is bounded for all x and n, but that

$$\sin x + \sin 2x + \ldots + \sin nx$$

is unbounded for small x and large n.

17. Prove that, if every a_r is positive and $x = \pi/2(n+p)$, then

$$|a_{n+1}\sin(n+1)x+..+a_{n+p}\sin(n+p)x| > \frac{p}{n+p}B_n,$$

where B_n is the least term (or the lowest bound) of the infinite set

$$na_n,\ (n+1)a_{n+1},\dots.$$

If it is known only that a_n is positive and that $\sum a_n \sin nx$ converges throughout a neighbourhood of $x = 0$, show that $\lim na_n = 0$ is a necessary condition for uniform convergence over such a neighbourhood

18. If $\{a_n\}$ be a positive sequence converging monotonically to zero and if α, β be any positive numbers, show that the condition

$$na_n \to 0$$

is both necessary and sufficient for the uniform convergence over every interval of the series

$$\sum a_n \sin(n\alpha+\beta)x$$

19 If $a, b, c,$ be integers mutually interprime, and if from the series $\sum (1/n)\sin nx$ every term is removed in which n is a multiple of any of $a, b, c,$, prove that, except for special values of x, the series that is left converges to the constant sum

$$\frac{\pi}{2}\left(1-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)\left(1-\frac{1}{c}\right) .$$

What are these special values of x?

II

THE CONTINUOUS FUNCTION

1. Convergence over a continuous domain

In common practice *convergence to a limit* is first defined with reference to a sequence, that is to say, to a one-valued function $f(n)$ converging over the positive integral domain. We must now extend the idea to convergence over a continuous domain. Two possible cases arise.†

The most immediate extension is to the convergence of $f(x)$ as x diverges to infinity.

We say that

(1) *$f(x)$ converges to the limit A as x diverges continuously, if for every positive ϵ we can find some positive $X(\epsilon)$ such that*

$$|f(x)-A| < \epsilon$$

for every x such that $|x| > X$.

We may still write
$$f(x) \to A \quad \text{as} \quad x \to \infty,$$
and, if x is restricted to positive (negative) values only, we may mark this by writing
$$f(x) \to A \quad \text{as} \quad x \to +\infty \ (-\infty).$$

The second case is that in which x converges to some number ξ. We have then to consider values of $f(x)$ in the neighbourhood of $x = \xi$. From this consideration we may choose to include or to exclude the value at the limit-point itself. It simplifies the discussion to begin by excluding this value. We then say that

(2) *$f(x)$ converges to the limit A as x converges continuously to ξ from below, if for every positive ϵ we can find a positive $\delta(\epsilon, \xi)$ such that*

$$|f(x)-A| < \epsilon$$

throughout the half-open interval $(\xi-\delta, \xi($.

We may write‡
$$f(x) \to A, \quad \text{as} \quad x \to \xi-.$$
We similarly define convergence from above by consideration of half-open intervals $)\xi, \xi+\delta)$ and write

$$f(x) \to A \quad \text{as} \quad x \to \xi+.$$

If $f(x)$ converges to A as x converges to ξ, whether from below or from above, we say simply that $f(x)$ converges to A as x converges to ξ and write
$$f(x) \to A, \quad \text{as} \quad x \to \xi;$$

† I here limit consideration to a one-valued function of a single variable.

‡ '$\xi-0$, $\xi+0$' is the accepted notation, but I omit the '0' as otiose.

or, for special emphasis,

$$f(x) \to A, \quad \text{as} \quad x \to \xi \pm.$$

We might further contemplate a function whose domain of definition was neither a complete interval nor yet the positive integral domain, for instance, a function defined for rational values of the argument alone. We could extend the idea of convergence to such a function in the spirit of the definitions (1), (2) above; we have merely to read 'every x in the domain of definition' for 'every x'. We shall not in general be concerned with such discontinuous domains of definition, but we need similar precautions with an elementary function that fails to be defined at isolated points through indeterminacy of its defining formula, e.g. if $f(x)$ takes the form $0/0$. In all such cases by 'every x' we shall imply 'every x at which the function is defined'.

2. Convergence over a restricted domain

Now, if the inequality $$|f(x)-A| < \epsilon$$

holds over any domain, it holds, of course, over any part of that domain. Hence, if we know that $f(x) \to A$ as $x \to \xi$ continuously, we know also that $f(x) \to A$ as $x \to \xi$ through any more restricted domain.

In particular

(3) *If $f(x) \to A$ as $x \to \xi$, and if $\{x_n\}$ is any sequence converging to ξ, then*

$$f(x_n) \to A \quad \text{as} \quad n \to \infty.$$

For, by (2), $$|f(x_n)-A| < \epsilon, \quad \text{if} \quad |x_n-\xi| < \delta(\epsilon, \xi).$$

The second inequality is secured, if $n >$ some $N(\delta)$, since $x_n \to \xi$.

I shall call the sequence $\{x_n\}$ a 'sequence of approach' to ξ and say that 'continuous' convergence secures convergence along any 'sequence of approach'. Conversely,

(4) *If $f(x)$ does not converge to A as x converges continuously to ξ, we can construct a sequence $\{x_n\}$ converging to ξ such that $f(x_n)$ does not converge to A as $n \to \infty$.*

We contradict (2) by interchanging 'some' and 'every' and reversing the central inequality. Hence, supposing non-convergence to A at ξ, then for some positive ϵ and every δ we can find some $x(\delta, \epsilon)$ in the interval $(\xi-\delta, \xi+\delta)$ such that

$$|f(x)-A| > \epsilon.$$

Fix this ϵ and choose a positive sequence $\{\delta_n\}$ converging to zero.

With each δ_n associate its $x_n(\delta_n, \epsilon)$. Then $x_n \to \xi$, since x_n lies in the interval $(\xi-\delta_n, \xi+\delta_n)$ and $\delta_n \to 0$. But by the property of x_n,

$$|f(x_n)-A| > \epsilon.$$

Hence $f(x_n)$ does not converge to A as $n \to \infty$. It follows from (4) that

(5) *If $f(x_n)$ converges to A for every sequence $\{x_n\}$ converging to ξ, then $f(x)$ converges to A as x converges continuously to ξ.*

These results are still true, if we separate 'convergence from above' and 'convergence from below', and they apply as well to x diverging.

So far we have spoken of convergence to a stated limit A. The 'general' condition of convergence, i.e. the condition of convergence in which the limit is unspecified follows the similar condition for a sequence:

(6) *The necessary and sufficient condition that $f(x)$ converge as $x \to \xi-$ is that for every positive ϵ we can find a positive $\delta(\epsilon, \xi)$ such that*

$$|f(x_1)-f(x_2)| < \epsilon$$

at all points x_1, x_2 in the half-open interval $(\xi-\delta, \xi($.

The condition is certainly necessary, for by (2), if δ is $\delta(\frac{1}{2}\epsilon, \xi)$, and x_1, x_2 lie in the interval $(\xi-\delta, \xi($, then

$$|f(x_1)-A| < \tfrac{1}{2}\epsilon, \qquad |f(x_2)-A| < \tfrac{1}{2}\epsilon,$$

and so
$$|f(x_1)-f(x_2)| < \epsilon.$$

Conversely, if (6) hold, let $\{x_n\}$ be any sequence converging to ξ from below. Then given ϵ we can find $N(\delta)$ such that $\xi-x_n < \delta(\epsilon, \xi)$, if $n > N$, and so x_n, x_{n+p} both lie in the interval $(\xi-\delta, \xi($. Hence, by (6),

$$|f(x_n)-f(x_{n+p})| < \epsilon,$$

if $n > N$. This secures the convergence of the sequence $\{f(x_n)\}$ to *some* limit A.

Suppose, if possible, that two such sequences $\{f(x_n)\}$, $\{f(x'_n)\}$ converge to different limits A, A'. Choose $\epsilon < \frac{1}{3}|A-A'|$. Then for the first sequence we can choose $n >$ some $N(\epsilon)$ such that

$$|f(x_n)-A| < \epsilon \tag{7}$$

and
$$\xi-x_n < \delta(\epsilon, \xi).$$

So for the second sequence we can choose $n >$ some $N'(\epsilon)$ such that

$$|f(x'_n)-A'| < \epsilon \tag{8}$$

and
$$\xi-x'_n < \delta(\epsilon, \xi).$$

Since both x_n, x'_n lie in $(\xi-\delta, \xi($,

$$|f(x_n)-f(x'_n)| < \epsilon. \tag{9}$$

From (7), (8), (9) we deduce that

$$|A-A'| < 3\epsilon,$$

which contradicts the choice of ϵ. Hence every sequence $\{f(x_n)\}$ converges to the same limit A, and therefore, by (5), $f(x)$ converges continuously to A from below.

We can state and establish similar general conditions of convergence for x converging to $\xi+$ or diverging.

If as x converges to ξ or diverges, $1/f(x)$ converges to zero, we say that $f(x)$ itself diverges. As with a sequence, the conditions of (6) are not satisfied by a diverging $f(x)$: divergence is a special case of non-convergence.

3. Continuity

Let us now think only of x converging to ξ, and let us bring into the discussion $f(\xi)$, the value at the limit-point.

Suppose that $f(\xi)$ converge as $x \to \xi-$. Write

$$f(x) \to f(\xi-).$$

Similarly, suppose that, as $x \to \xi+$,

$$f(x) \to f(\xi+).$$

Then with the point $x = \xi$ are associated the three numbers

$$f(\xi-), \qquad f(\xi+), \qquad f(\xi),$$

i.e. the limit from below, the limit from above, and the value at the point. If $f(\xi-) = f(\xi+)$ we may write this value $f(\xi\pm)$ and call it the 'limit at the point'. If finally the limit at the point and the value at the point are the same, we say that the function is 'continuous' at the point. In other words,

(10) *$f(x)$ is said to be continuous at $x = \xi$, if $f(x) \to f(\xi)$ as $x \to \xi$ both from above and below.*

We may say that at a point of continuity

$$\lim f(x) = f(\lim x).$$

The distinction between convergence and continuity at a point is then this: the function is convergent, if it has a limit at the point; it is continuous, if it has both a limit and a value at the point and these are equal. The distinction is not generally important for elementary functions, but it may be convenient to mark it at a point of non-definition of a function.

Thus, strictly, the function $(\sin x)/x$ is undefined at $x = 0$. It converges there to the value unity. Hence for precision we should say

that the function is 'convergent' not 'continuous' at the point. But in common practice we are accustomed, either explicitly or by implication, to add in this limiting value to the values of the function. In effect we define

$$f(x) = (\sin x)/x \qquad (x \neq 0)$$

$$f(0) = 1.$$

By this device we extend the function convergent at $x = 0$ into a function continuous at that point.

If, in our definition of convergence, we replace the limiting value A by $f(\xi)$ and at the same time close the intervals $(\xi \pm \delta, \xi)$, we have the definition of continuity:

(11) *$f(x)$ is continuous at $x = \xi$, if for every ϵ we can find $\delta(\epsilon, \xi)$ such that*

$$|f(x) - f(\xi)| < \epsilon$$

throughout the interval $(\xi - \delta, \xi + \delta)$.

The general condition of convergence (6), if we close the intervals $(\xi \pm \delta, \xi)$, becomes a mere variant of (11). For, if we know that

$$|f(x_1) - f(x_2)| < \epsilon$$

throughout an interval $(\xi - \delta, \xi + \delta)$, where $\delta = \delta(\epsilon, \xi)$, we may take x_2 to be ξ itself and we at once reproduce (11).

If $f(x)$ is defined only throughout some interval $a \leqslant x \leqslant b$, then convergence of $f(x)$ at $x = a$ can only mean convergence as x tends to a from above. If, precisely, $f(x) \to f(a)$ as $x \to a+$, we shall say that $f(x)$ is continuous at $x = a$. So, generally, by continuity at an end-point of the domain of definition we must understand continuity as we approach the end-point from within the domain of definition, and we must make corresponding modifications in the analytical conditions for continuity at such an end-point.

4. Elementary properties of a function at a point of continuity

The analytical conditions for continuity of a function differ from those for convergence of a sequence only in the precise stipulation of the limit and in the nature of the domain through which the fundamental inequality is to be satisfied; the inequality itself does not differ in character. It is not therefore difficult to see that we may take over certain elementary propositions from the Theory of Convergence. Thus

(12) *The sum, the difference, and the product of functions, continuous at a given point, are themselves continuous at the point.*

It follows by repeated application of (12) that

(13) *A rational, integral algebraic function of functions continuous at a given point is itself continuous at that point.*

Evidently x regarded as a function of itself is continuous at every point.† Hence, by (13),

(14) *A rational, integral algebraic function is everywhere continuous.*

If a function is continuous at all points of a region, we say that it is continuous *throughout* the region. A function continuous throughout its domain of definition we may call briefly a 'continuous function'. Hence by (13), (14) we may say that

(15) *All rational, integral algebraic functions are continuous functions; all rational, integral algebraic functions of continuous functions are themselves continuous functions.*

More generally we can prove that

(16) *A continuous function of a continuous function is itself a continuous function.*

For let $z = z(y)$ and $y = y(x)$ be the two continuous functions.

Then we secure $\quad |z(y)-z(\eta)| < \epsilon$

by taking $\quad |y-\eta| <$ some $\delta(\epsilon, \eta)$,

i.e. by taking $\quad |y(x)-y(\xi)| < \delta$.

But, since $y(x)$ is continuous, we secure this second inequality if

$$|x-\xi| < \text{some } \gamma(\delta, \xi).$$

From the Theory of Convergence we may take over this further proposition:

(17) *The ratio of two functions continuous at a point that is not a zero of the denominator is itself continuous at the point.*

Hence by (15) we have that

(18) *A rational algebraic function is everywhere continuous save at the zeros of the denominator.*

At the zeros of the denominator the function is divergent.

5. Continuity of x^p, a^x, $\log x$

When a is negative, a^x is real only for certain rational values of x; when a is positive, a^x is negative only for certain other rational values of x. Hence, as a real function defined throughout a continuous domain, a^x exists only for positive values of the base a and is itself positive. The function so conditioned is then also one-valued.

† For $x \to \xi$ as $x \to \xi$.

It is convenient to make similar stipulations for the function x^p, regarding it as a one-valued, positive function of a positive argument; when $p > 0$, we may extend the domain to $x = 0$, writing $x^p = 0$, when $x = 0$. The function x^p so defined satisfies the 'exponential inequality'†

$$p(x-\xi)x^{p-1} \gtrless x^p - \xi^p \gtrless p(x-\xi)\xi^{p-1}. \tag{19}$$

Hence $|x^p - \xi^p|$ lies between

$$|p(x-\xi)|x^{p-1} \quad \text{and} \quad |p(x-\xi)|\xi^{p-1},$$

and so between

$$|p(x-\xi)|(\tfrac{1}{2}\xi)^{p-1} \quad \text{and} \quad |p(x-\xi)|(\tfrac{3}{2}\xi)^{p-1},$$

if x lies in the interval $(\frac{1}{2}\xi, \frac{3}{2}\xi)$. Accordingly, by taking

$$\delta < \min\{\tfrac{1}{2}\xi,\ \epsilon|p|^{-1}(\tfrac{1}{2}\xi)^{-p+1},\ \epsilon|p|^{-1}(\tfrac{3}{2}\xi)^{-p+1}\},$$

we secure
$$|x^p - \xi^p| < \epsilon$$
throughout the interval $(\xi-\delta, \xi+\delta)$.

The analysis fails, if $\xi = 0$, but then, if $p > 0$, we have

$$|x^p - \xi^p| = x^p < \epsilon, \quad \text{if} \quad x < \epsilon^{1/p}.$$

If $p = 0$, x^p is constant for every positive x and is therefore convergent at $x = 0$. We may make it continuous there, if we define $x^0 = 1$ at $x = 0$. If $p < 0$, x^p is divergent at $x = 0$. Thus

(20) *x^p is everywhere continuous, save at $x = 0$ when $p < 0$.*

To discuss a^x rewrite the exponential inequality as

$$(x-\xi)(a-1)a^{x-\xi-1} \gtrless a^{x-\xi} - 1 \gtrless (x-\xi)(a-1).$$

Then
$$|a^x - a^\xi| = a^\xi|a^{x-\xi} - 1|$$
and so lies between

$$|(x-\xi)(a-1)|a^{\xi-2} \quad \text{and} \quad |(x-\xi)(a-1)|a^{\xi},$$

if x lies in the interval $(\xi-1, \xi+1)$. Accordingly, by taking

$$\delta < \min(1,\ \epsilon|a-1|^{-1}a^{2-\xi},\ \epsilon|a-1|^{-1}a^{-\xi}),$$

we secure
$$|a^x - a^\xi| < \epsilon$$
throughout the interval $(\xi-\delta, \xi+\delta)$. The analysis fails if $a = 1, 0$. But the function 1^x is everywhere constant and therefore everywhere continuous. The function 0^x is undefined for negative x, but is constant and therefore continuous for positive x In particular 0^x converges to 0 as x converges to $0+$, and we therefore make 0^x continuous at $x = 0$ by defining $0^x = 0$ when $x = 0$.‡ Thus, in sum,

(21) *a^x is a continuous function for $a \geqslant 0$.*

† See, for instance, Chrystal, *Algebra*, 2 (1900), 45 (Corollary).

‡ We may remark that to secure the continuity of both x^0 and 0^x at $x = 0$ we have been led into the inconsistent definitions $0^0 = 1$, $0^0 = 0$.

For the logarithmic function, it is enough to consider only the base e, since $\log_a x$ is a constant multiple of $\log_e x$. The function is defined only for positive x and we have the characteristic inequality†

$$c-1 \geqslant \log c \geqslant (c-1)/c \qquad (c > 0).$$

Thus $|\log(x/\xi)|$ lies between

$$|x-\xi|/\xi, \qquad |x-\xi|/x.$$

Hence, by taking $$\delta < \epsilon\xi/(1+\epsilon),$$

we have $$|\log x - \log \xi| < \epsilon$$

throughout the interval $(\xi-\delta, \xi+\delta)$. At $x = 0$ the function is undefined, but, since $$|\log x| > 1/\epsilon, \quad \text{if} \quad x < \exp(-1/\epsilon),$$

we see that $\log x$ diverges as $x \to 0+$. Thus

(22) $\log x$ *is continuous for every positive* x.

6. Continuity of the circular and the hyperbolic functions

For the circular functions the fundamental inequality‡ is

$$|\sin x| < |x|.$$

We then have

$$|\sin x - \sin \xi| = 2\left|\sin\frac{x-\xi}{2}\cos\frac{x+\xi}{2}\right| < |x-\xi|.$$

Hence $|\sin x - \sin \xi| < \epsilon$ throughout the interval $(\xi-\epsilon, \xi+\epsilon)$. The sine is therefore a continuous function. Since $\cos x = \sin(\frac{1}{2}\pi - x)$, we see from (16) that the cosine is also a continuous function.

It follows from (17) that $\tan x$, $\sec x$ are continuous except at $x = (n+\frac{1}{2})\pi$, and $\cot x$, $\operatorname{cosec} x$ continuous except at $x = n\pi$, where n is an integer or zero.

The inverse circular functions arise as many-valued functions, but for each direct circular function there is a pair of adjacent quadrants in which the function runs through all its possible values once and once only. The corresponding inverse function, restricted by explicit definition to this pair of quadrants, becomes a single-valued function. It is the familiar convention of 'principal values'.

In practice we restrict $\sin^{-1}x$, $\tan^{-1}x$, $\operatorname{cosec}^{-1}x$ to the interval $)-\frac{1}{2}\pi, \frac{1}{2}\pi)$ and therefore their co-functions $\cos^{-1}x$, $\cot^{-1}x$, $\sec^{-1}x$, i.e. $\frac{1}{2}\pi - \sin^{-1}x$, etc., to the corresponding interval $(0, \pi($.

In the restricted interval we have the inequality§

$$|\tan^{-1}x| < |x|.$$

† See also Chrystal, loc. cit., 80 (Corollary 5).

‡ See, for instance, Hobson, *Plane Trigonometry* (1897), 122.

§ See also Hobson, ibid.

Now, if x, ξ have the same sign or $\xi = 0$,†

$$\tan^{-1}x - \tan^{-1}\xi = \tan^{-1}\frac{x-\xi}{1+x\xi},$$

and so
$$|\tan^{-1}x - \tan^{-1}\xi| < \left|\frac{x-\xi}{1+x\xi}\right| \leqslant |x-\xi|,$$

since $x\xi$ is positive or zero. Hence $|\tan^{-1}x - \tan^{-1}\xi| < \epsilon$ throughout the interval $(\xi-\delta, \xi+\delta)$, if $\delta < \min(\epsilon, |\xi|)$. Thus $\tan^{-1}x$ is a continuous function and therefore also its co-function $\cot^{-1}x$. Moreover, $\tan^{-1}x$ converges as x diverges. For suppose x positive. Then

$$\tfrac{1}{2}\pi - \tan^{-1}x = \cot^{-1}x = \tan^{-1}(x^{-1}).$$

Hence $\tan^{-1}x \to \frac{1}{2}\pi$, as $x \to +\infty$; similarly $\tan^{-1}x \to -\frac{1}{2}\pi$, as $x \to -\infty$.

The inverse secant does not exist as a real function in the open interval $)-1, 1($. Elsewhere it is defined in terms of the inverse tangent, having regard to the convention of 'principal values', by the formulae

$$\begin{aligned} \sec^{-1}x &= \tan^{-1}\sqrt{(x^2-1)} \qquad && (x \geqslant 1), \\ &= \pi - \tan^{-1}\sqrt{(x^2-1)} \qquad && (x \leqslant -1). \end{aligned}$$

Since $\sqrt{(x^2-1)}$ is a continuous function, so also is $\sec^{-1}x$, and therefore, too, $\operatorname{cosec}^{-1}x$. Moreover, we find that

$$\sec^{-1}x \to \tfrac{1}{2}\pi, \quad \operatorname{cosec}^{-1}x \to 0, \quad \text{as } x \to \pm\infty.$$

The inverse sine is defined only over the closed interval $(-1, 1)$; we have, in fact,
$$\sin^{-1}x = \operatorname{cosec}^{-1}(x^{-1}).$$

Hence $\sin^{-1}x$ is continuous except possibly at $x = 0$. But, since $\operatorname{cosec}^{-1}x \to 0$ as $x \to \pm\infty$, we see that $\sin^{-1}x \to \sin^{-1}0$, as $x \to 0$, and thus that $\sin^{-1}x$ is continuous throughout its domain of definition; so therefore is $\cos^{-1}x$. In sum

(23) *Of the circular functions*, $\sin x$, $\cos x$ *are continuous everywhere*; $\tan x$, $\sec x$ *except at* $x = (n+\frac{1}{2})\pi$; $\cot x$, $\operatorname{cosec} x$ *except at* $x = n\pi$, *where* n *is integral or zero.*

Of the inverse circular functions, $\sin^{-1}x$, $\cos^{-1}x$, *are continuous, if* $|x| \leqslant 1$; $\tan^{-1}x$, $\cot^{-1}x$ *everywhere*; $\sec^{-1}x$, $\operatorname{cosec}^{-1}x$, *if* $|x| \geqslant 1$.

The hyperbolic functions $\sinh x$, $\cosh x$, $\tanh x$, $\coth x$, $\operatorname{sech} x$, $\operatorname{cosech} x$ are all defined (or definable) as rational algebraic functions of e^x. Since the latter is everywhere continuous, the hyperbolic functions are continuous except where they become infinite. We find then that

$$\sinh x, \ \cosh x, \ \tanh x, \ \operatorname{sech} x$$

† This stipulation is necessary, since otherwise $\tan^{-1}x - \tan^{-1}\xi$ might numerically exceed $\frac{1}{2}\pi$ and so could not be the 'principal value' of an inverse tangent.

are everywhere continuous;

$$\coth x, \ \operatorname{cosech} x$$

are continuous everywhere except at $x = 0$.

The inverse hyperbolic functions can be expressed as logarithms of algebraic functions, e.g.

$$\sinh^{-1}x = \log\{x+\sqrt{(x^2+1)}\}, \qquad \cosh^{-1}x = \log\{x\pm\sqrt{(x^2-1)}\}.$$

They are one-valued with the exception of $\cosh^{-1}x$, $\operatorname{sech}^{-1}x$, and these can be made one-valued by taking their positive values as 'principal' values. With the exception of $\sinh x$, the hyperbolic functions cannot assume all real values for real values of the argument. The domains of definition of the associated inverse function are therefore correspondingly restricted. Bearing these facts in mind we find that

$$\sinh^{-1}x \text{ is continuous everywhere,}$$

$$\operatorname{cosech}^{-1}x, \text{ if } x \neq 0, \qquad \cosh^{-1}x, \text{ if } x \geqslant 1,$$

$$\operatorname{sech}^{-1}x, \text{ if } 0 < x \leqslant 1, \qquad \tanh^{-1}x, \text{ if } |x| < 1,$$

$$\coth^{-1}x, \text{ if } |x| > 1.$$

It is to be noted that all the foregoing elementary functions are continuous wherever they are defined. Continuity fails only by divergence of the function at an end-point of an open interval of definition.

7. Continuity of limit-functions and sum-functions

It would have been possible to discuss the continuity of the elementary transcendental functions from their expansions in infinite series, and it is indeed usually necessary to approach the higher transcendental functions in some such way. We are thus led to study the continuity of functions defined as sum-functions of infinite series or, more generally, as limit-functions of sequences.

An example will show that a sequence of continuous functions does not necessarily define a continuous limit-function, i.e. continuity need not survive passage to the limit. Consider the sequence

$$s(n,x) \equiv (1+x^2)^{-n}. \tag{24}$$

Then, if $x \neq 0$, $s(n,x) \to 0$ as $n \to \infty$, but $s(n,0) = 1$, and so $s(n,0) \to 1$ as $n \to \infty$. Thus the limit-function $s(x)$ is given by

$$s(x) = 0 \quad (x \neq 0), \qquad s(0) = 1.$$

There is therefore discontinuity at $x = 0$.

A *sufficient* condition for continuity of the limit-function is afforded

by the uniform convergence of the sequence. Precisely,

(25) *If for every n the function $s(n,x)$ is continuous at $x = \xi$, and if $s(n,x)$ converges to $s(x)$ uniformly over some neighbourhood of ξ, then $s(x)$ is continuous at $x = \xi$.*

Suppose the convergence to be uniform over the interval $(\xi-\delta, \xi+\delta)$. Then

$$|s(n,x)-s(x)| < \tfrac{1}{3}\epsilon$$

for every $n >$ some $N(\epsilon)$ and every x in the interval. In particular

$$|s(n,\xi)-s(\xi)| < \tfrac{1}{3}\epsilon.$$

Fix one such n. Then from the continuity of $s(n,x)$ we have

$$|s(n,x)-s(n,\xi)| < \tfrac{1}{3}\epsilon,$$

if $|x-\xi| <$ some $\delta_1(n,\epsilon)$. By combining the three inequalities we have

$$|s(x)-s(\xi)| < \epsilon,$$

if $|x-\xi| < \min(\delta, \delta_1)$. The limit-function is therefore continuous at $x = \xi$.

It follows that the sequence (24) cannot converge uniformly in the neighbourhood of $x = \xi$. This is evident directly, for

$$(1+x^2)^{-n} < \epsilon, \quad \text{if} \quad n > (\log 1/\epsilon)/\log(1+x^2),$$

and, for a fixed ϵ,

$$(\log 1/\epsilon)/\log(1+x^2) \to \infty, \quad \text{as} \quad x \to 0.$$

A more notorious example of non-uniformity of convergence associated with discontinuity in the limit-function is afforded by the elementary Fourier's series

$$s(x) = \sin x + \tfrac{1}{2}\sin 2x + \tfrac{1}{3}\sin 3x + \dots \text{ to } \infty. \tag{26}$$

By direct substitution it is evident that

$$s(0) = 0.$$

If x lies in the interval $)0, 2\pi($, it is known† that the sum-to-infinity is

$$s(x) = \tfrac{1}{2}(\pi - x),$$

and hence $$s(0+) = \tfrac{1}{2}\pi.$$

Evidently $s(-x) = -s(x)$ and so, if x lies in the interval $)-2\pi, 0($,

$$s(x) = -\tfrac{1}{2}(\pi + x),$$

and hence $$s(0-) = -\tfrac{1}{2}\pi.$$

Thus at $x = 0$ we have

$$s(0-) = -\tfrac{1}{2}\pi, \qquad s(0) = 0, \qquad s(0+) = \tfrac{1}{2}\pi.$$

There is thus convergence from above and from below, but not continuity, since neither limit equals the value at the point.

† This is the result proved subsequently in chapter XIV § 4 (20) and already quoted in somewhat different form as chapter I (6).

Since the series is periodic with period 2π, there is, in the same way, discontinuity at the points $x = 2m\pi$, where m is integral. We have already seen† that this series is non-uniformly convergent in the neighbourhood of points $x = 2m\pi$.

If we apply (25) to a power-series, we have at once that

(27) *The sum-function of a power-series is continuous wherever the power-series converges.*

For, by chapter I § 11 (15), the power-series $s(n,x)$ converges uniformly over $(0,\xi)$, if it converges at $x = \xi$. Suppose ξ positive. Then, by (25), $s(x) \to s(\xi)$, as $x \to \xi-$. This is sufficient to establish continuity at $x = \xi$, if it is an end-point of the interval of convergence, since $s(x)$ is not defined for $\xi+$.

If ξ is not an end-point, then $s(n,x)$ also converges at some point $\xi+h$, where h is positive, and hence converges uniformly over $(0,\xi+h)$. By (25) this secures continuity at $x = \xi$.

Thus a power-series is continuous wherever it is defined. In other words, we may say that

(28) *The sum-function of a power-series is a continuous function.*

This result evidently embraces the continuity of the elementary functions e^x, $\log(1+x)$, $\sin x$, $\cos x$.

On the other hand, it is not true, conversely, that 'if the sum-function $s(x)$ converges, as $x \to \xi$, then the power-series $s(n,x)$ converges at $x = \xi$'. This is shown by a simple example

$$s(n,x) \equiv 1-x+x^2-\ldots+(-x)^n.$$

Here $$s(x) = (1+x)^{-1} \to \tfrac{1}{2}, \quad \text{as} \quad x \to 1,$$

but $$s(n,1) = 1-1+1-1+\ldots,$$

an oscillating series.

More generally, we know that the expansion of $(1+x)^p$ by the binomial theorem converges at $x = 1$, if and only if $p > -1$, whereas the sum-function $(1+x)^p$ converges as $x \to 1$ for every p.

Tauber and others have considered conditions sufficient for the converse in question to be true.‡

We can here establish a proposition that was anticipated in our discussion of the uniform convergence of a power-series:

(29) *An interval of uniform convergence of a sequence of continuous functions is a closed interval.*

† Chapter I § 13.

‡ See, in particular, Hardy and Littlewood, *Proc. London Math. Soc.* (2) 30 (1930), 23–37, where important references are given.

For suppose that $s(n,x)$ converge uniformly over the open interval $h < x < k$. Then given ϵ we can find $N(\epsilon)$ such that

$$|s_{n+p}(x) - s_n(x)| < \epsilon,$$

if $n > N$ and $h < x < k$. Since N is independent of x, we can fix ϵ, n, p and take the limit as $x \to k$. The functions are continuous at $x = k$ and so

$$|s_{n+p}(k) - s_n(k)| \leqslant \epsilon.$$

Hence
$$|s_{n+p}(x) - s_n(x)| < 2\epsilon,$$

if $n > N$ and $h < x \leqslant k$. We have thus closed the interval of uniform convergence at $x = k$ and we can similarly close it at $x = h$.

Hence, in particular, the interval of uniform convergence of a power-series is always closed and therefore cannot coincide with the interval of ordinary convergence, except when this latter interval is closed.

The example
$$s(n,x) = \{I(x)\}^n$$

of chapter I § 12 (22) does not contradict (29). It converges uniformly over the half-open interval $(1, 2($, which cannot be closed at $x = 2$, since $s(n, 2)$ diverges. The functions $s(n,x)$ are, however, discontinuous at $x = 2$.

8. Vanishingly uniform convergence

The condition of uniform convergence, though sufficient, is not necessary for continuity of the limit-function. This can be shown by an example

$$s(n,x) = \frac{nx}{1+n^2x^2}. \tag{30}$$

If $x \neq 0$, $s(n,x) \to 0$ and $s(n,0) = 0$. Hence the limit-function $s(x)$ is always zero and, in particular, is continuous at $x = 0$. In the neighbourhood of $x = 0$ we can write $x = 1/n$ for arbitrarily large n; this gives $s(n,x) = \frac{1}{2}$. Hence, by chapter I § 12 (18), the convergence is not uniform in the neighbourhood of $x = 0$.

Thus uniform convergence is not necessary for continuity of the limit-function. An exact condition can be given as follows:

(31) *The necessary and sufficient condition that the convergent sequence of functions $\{s(n,x)\}$, all continuous at $x = \xi$, define a limit-function $s(x)$, itself continuous at $x = \xi$, is that, given any positive ϵ and N,*

$$|s(n,x) - s(x)| < \epsilon$$

for some $n(\epsilon, \xi) > N$ and every x in some interval $|x - \xi| < \delta(\epsilon, \xi, n)$.

The condition is necessary, for, given ϵ and N, then by the continuity of $s(x)$ we can find $\delta(\epsilon, \xi)$ such that

$$|s(x) - s(\xi)| < \tfrac{1}{3}\epsilon, \quad \text{if} \quad |x - \xi| < \delta(\epsilon, \xi). \tag{32}$$

Since $s(n,\xi) \to s(\xi)$, we can find $N_1(\epsilon,\xi)$ such that

$$|s(n,\xi)-s(\xi)| < \tfrac{1}{3}\epsilon, \quad \text{if} \quad n > N_1. \tag{33}$$

Fix any $n > \max(N, N_1)$. Then, since $s(n,x)$ is continuous at $x = \xi$, we can find $\delta'(\epsilon,\xi,n)$ such that

$$|s(n,x)-s(n,\xi)| < \tfrac{1}{3}\epsilon, \quad \text{if} \quad |x-\xi| < \delta'. \tag{34}$$

Thus, if δ_0 is the smaller of δ, δ', we have, by combining (32), (33), (34),

$$|s(n,x)-s(x)| < \epsilon, \quad \text{if} \quad |x-\xi| < \delta_0,$$

i.e. throughout the open interval $(\xi-\delta_0, \xi+\delta_0)$.

The condition is also sufficient. For, given ϵ, we can still find $N_1(\epsilon,\xi)$ to secure (33). Hence by the given condition we can find $n > N_1$ and $\delta(\epsilon,\xi,n)$ such that

$$|s(n,x)-s(x)| < \tfrac{1}{3}\epsilon, \quad \text{if} \quad |x-\xi| < \delta(\epsilon,\xi,n). \tag{35}$$

Lastly, with this same n, we can still find $\delta'(\epsilon,\xi,n)$ to secure (34). Hence, if $\delta_0(\epsilon,\xi,n)$ is the smaller of δ, δ', we have, by combining (33), (34), (35),

$$|s(x)-s(\xi)| < \epsilon, \quad \text{if} \quad |x-\xi| < \delta_0(\epsilon,\xi,n). \tag{36}$$

Since n no longer appears in the equality (36), we can write $\delta_0(\epsilon,\xi)$ for $\delta_0(\epsilon,\xi,n)$ and so the continuity of $s(x)$ at $x = \xi$ is established.

To the convergence prescribed in (31) we may give the name of 'vanishingly uniform convergence',† for, in general, $\delta(\epsilon,\xi,n) \to 0$ as $\epsilon \to 0$, $n \to \infty$. For, if $\delta(\epsilon,\xi,n)$ remains greater than some fixed δ_0 as $\epsilon \to 0$, $n \to \infty$, then the convergence is evidently uniform over the interval $(\xi-\delta_0, \xi+\delta_0)$. The distinction between 'uniform convergence' and 'vanishingly uniform convergence' lies essentially in the order of choice of ϵ, δ, n. With uniform convergence we can declare δ at the outset, i.e. the order of choice is

$$\delta,\ \epsilon,\ n.$$

For vanishingly uniform convergence the order of choice prescribed in (31) is

$$\epsilon,\ n,\ \delta.$$

Since, in (31), $\delta(\epsilon,\xi,n)$ may diminish indefinitely as $n \to \infty$, the condition

$$|s(n,x)-s(x)| < \epsilon, \quad \text{if} \quad |x-\xi| < \delta(\epsilon,\xi,n), \quad n > N$$

cannot of itself secure the convergence of $s(n,x)$ to $s(x)$, except possibly at $x = \xi$. But even here it is insufficient, since the inequality holds for some but not every n greater than N.

† For this and other varieties of uniform convergence see Bromwich, *Infinite Series* (1926), 138–40 (§ 49.1).

The form of condition in (31) presupposes that we know the limit-function $s(x)$, in which case it is presumably much easier to apply the test for continuity direct to $s(x)$ itself. A form of condition can be given in which $s(x)$ does not appear, corresponding to the 'general' conditions of convergence and of uniform convergence. Similar conditions can also be obtained for the *convergence* (as opposed to the *continuity*) of $s(x)$ as $x \to \xi+$, say. I state these conditions without proof:

(37) *The convergence of the sequence $s(n,x)$ is vanishingly uniform near $x = \xi$, if and only if, given ϵ and N,*

$$|s(n,x)-s(n+p,x)| < \epsilon$$

for some $n(\epsilon,\xi) > N$, every x in some interval $|x-\xi| < \delta(\epsilon,\xi,n)$, and every $p >$ some $p_0(\epsilon,\xi,n,x)$.

We can also replace

'every $p >$ some $p_0(\epsilon,\xi,n,x)$' by 'some $p(\epsilon,\xi,n,x) >$ a given P'

in either the necessary or the sufficient condition. The 'order of choice' is ϵ, n, δ, x, p.

For convergence, as $x \to \xi+$, we rule out of consideration the values at ξ both of $s(x)$ and of $s(n,x)$ and state the conditions:

(38) *If the sequence of functions $\{s(n,x)\}$, which all converge as $x \to \xi+$, also converge, as $n \to \infty$, to $s(x)$, this limit-function $s(x)$ also converges as $x \to \xi+$, if and only if, given ϵ and N,*

$$|s(n,x)-s(n,x')-s(x)+s(x')| < \epsilon$$

for some $n(\epsilon,\xi) > N$, and every x, x' in $)\xi,\xi+\delta)$ where $\delta = \delta(\epsilon,\xi,n)$; and

(39) *If the sequence of functions $\{s(n,x)\}$, which all converge as $x \to \xi+$, also converge as $n \to \infty$, then the limit-function also converges as $x \to \xi+$, if and only if, given ϵ and N,*

$$|s(n,x)-s(n,x')-s(n+p,x)+s(n+p,x')| < \epsilon$$

for some $n(\epsilon,\xi) > N$, every x, x' in $)\xi,\xi+\delta)$ where $\delta = \delta(\epsilon,\xi,n)$, and every $p >$ some $p_0(\epsilon,\xi,n,x,x')$.

Here again we may replace 'every $p >$ some $p_0(\epsilon,\xi,n,x,x')$' by 'some $p(\epsilon,\xi,n,x,x') >$ a given P' in either the necessary or the sufficient condition. The order of choice is ϵ, n, δ, x, x', p.

It may be added that the vanishingly uniform convergence of $s(n,x)$ to $s(x)$ in an interval $)\xi,\xi+\delta)$ is sufficient but not necessary for the

convergence of $s(x)$ as $x \to \xi+$. We can show that it is not necessary by an example

$$s(n,x) \equiv (1+nx)^{-1}. \tag{40}$$

If $x \neq 0$, then $s(x) = 0$ and

$$|s(n,x)-s(x)| = (1+nx)^{-1} > \epsilon, \quad \text{where} \quad 0 < \epsilon < 1,$$

if

$$x < (1-\epsilon)/\epsilon n,$$

which is incompatible with the condition of vanishingly uniform convergence, namely,

$$|s(n,x)-s(x)| < \epsilon,$$

if $x <$ some $\delta(n)$.

9. Types of discontinuity

A point at which a function is not continuous is called a 'discontinuity' or better a 'point of discontinuity' of the function. We may classify discontinuities into several types.

(i) If $f(\xi+)$, $f(\xi-)$, $f(\xi)$ all exist but are not all equal, i.e. if there is convergence without continuity, we call ξ a 'point of simple discontinuity' or we may say simply that there is a 'break' in the function at ξ. The greatest of

$$|f(\xi+)-f(\xi)|, \qquad |f(\xi-)-f(\xi)|, \qquad |f(\xi+)-f(\xi-)|$$

is the 'measure' of the discontinuity.

Such a discontinuity is characteristic of the sum-function of a Fourier's series $\sum a_n \sin nx$. For, as we have seen, with

$$s(x) \equiv \sin x+\tfrac{1}{2}\sin 2x+\tfrac{1}{3}\sin 3x+\ldots \text{ to } \infty$$

we have

$$s(0-) = -\tfrac{1}{2}\pi, \qquad s(0) = 0, \qquad s(0+) = \tfrac{1}{2}\pi.$$

There is a break in the function (and in its graph) at $x = 0$, and generally at $x = 2m\pi$, the measure of the break being π.

Again, if $I(x)$ denote the greatest integer in x, then, at any integer point $x = n$,

$$I(n-) = n-1, \qquad I(n) = n, \qquad I(n+) = n.$$

There is at every integer point a break in the function of measure unity.

(ii) If $1/f(x)$ is continuous but zero at $x = \xi$, so that $f(\xi)$ is undefined and $f(x)$ diverges as $x \to \xi$, we may call ξ a 'simple infinity' or a 'pole'. This is the only type of discontinuity which occurs with algebraic functions. It occurs at the zeros of the denominator of fractional functions. Thus $x = 0$ is a pole of

$$1/x, \qquad \cot x, \qquad \log|x|.$$

(iii) A discontinuity may arise which is a blend of the simple discontinuity and the simple infinity. Thus

$$\begin{aligned} \exp(x^{-1}) &\to \infty \qquad (x \to 0+), \\ &\to 0 \qquad (x \to 0-). \end{aligned}$$

It is better to regard such a discontinuity as a simple discontinuity whose measure is infinite.

(iv) If, as $x \to \xi-$ say, $f(x)$ neither converges nor diverges, we say that $f(x)$ 'oscillates' as $x \to \xi-$. It oscillates finitely or infinitely according as $f(x)$ is bounded or unbounded in some interval $(\xi-\delta, \xi)$, i.e. according as we can or cannot assign some $M(\delta)$ such that $|f(x)| < M$ throughout $(\xi-\delta, \xi)$.

If $f(x)$ oscillates, as $x \to \xi-$ or as $x \to \xi+$, we call ξ a 'point of oscillation'. It is a point of finite oscillation, if $f(x)$ is bounded in the neighbourhood of ξ; it is a point of infinite oscillation, if $f(x)$ is unbounded in the neighbourhood of ξ.

In illustration consider the following functions at $x = 0$:

$$\begin{aligned} f_1(x) &\equiv \sin(x^{-1}), \\ f_2(x) &\equiv x^{-1}\sin(x^{-1}), \\ f_3(x) &\equiv \exp(x^{-1})+\sin(x^{-1}). \end{aligned}$$

In $f_1(x)$, if x travels to the origin along the sequence of approach

$$\alpha^{-1}, (\alpha+2\pi)^{-1},..., (\alpha+2n\pi)^{-1},...,$$

where α is arbitrary, $\sin(x^{-1})$ has the constant value $\sin\alpha$, and hence, along this sequence of approach

$$\sin(x^{-1}) \to \sin\alpha.$$

Thus $f_1(x)$ oscillates as $x \to 0$, and, since $|f_1(x)| < 1$, the origin is a point of *finite oscillation* of $f_1(x)$.

In $f_2(x)$ the factor x^{-1} diverges as $x \to 0$. Hence $f_2(x)$ has the origin for a point of *infinite oscillation*.

Again $f_3(x) \to \infty$ as $x \to 0+$, since $\exp(x^{-1})$ diverges and so smothers the finite oscillations of $\sin(x^{-1})$. But, as $x \to 0-$, $\exp(x^{-1}) \to 0$ and the finite oscillations of $\sin(x^{-1})$ now predominate. Thus $f_3(x)$ oscillates finitely as $x \to 0-$. It is best to consider the origin a point of infinite oscillation of such a function as $f_3(x)$.

We may call a point ξ an 'infinity' of $f(x)$, if $f(x)$ is unbounded in every neighbourhood of ξ. Thus an infinity may be (i) a pole, (ii) a break of infinite measure, or (iii) a point of infinite oscillation.

WORKED EXAMPLE

Every number x lying in $(0, 1)$ is written as a decimal

$$x = 0{\cdot}a_1 a_2 a_3 \ldots .$$

The corresponding value of $f(x)$ is defined as the decimal

$$f(x) = 0{\cdot}b_1 b_2 b_3, \ldots ,$$

where $b_n = \phi_n(a_1, a_2, \ldots, a_m)$ and $m = m(n)$. Show that, in general, $f(x)$ is discontinuous for every value of x which can be expressed as a terminating decimal, but is continuous for every other value of x.†

In the first place we observe that every number which can be expressed as a terminating decimal can also be expressed as a recurring decimal. For instance

$$\tfrac{1}{2} = 0{\cdot}5 = 0{\cdot}4999 \ldots .$$

Any complete and 'monotonic' representation of the aggregate of real numbers by sequences of positive integers leads necessarily to alternative representations of certain numbers. For example, since the decimal representation is monotonic, every number $0{\cdot}4\ldots$ is less than every number $0{\cdot}5\ldots$. Moreover, $0{\cdot}4999\ldots$ is greatest of the numbers $0{\cdot}4\ldots$ and $0{\cdot}5000\ldots$ is the least of the numbers $0{\cdot}5\ldots$. Thus, if $0{\cdot}4999\ldots$ and $0{\cdot}5000\ldots$ represented distinct numbers, the numbers which lay between them would be unrepresented and the representation of real numbers would not be complete.

Suppose now that ξ has a unique representation $0{\cdot}a_1 a_2 a_3 \ldots$. It must have infinitely many digits a_n other than 9, the suffixes of these digits forming an infinite sequence $\{M_r\}$, say. For any such M_r, the representation of every x such that

$$0 \leqslant x - \xi < 10^{-M_r}$$

begins with the same $M_r - 1$ digits. Determine the greatest n_r such that

$$m(n) \leqslant M_r - 1 \qquad (n < n_r).$$

Then all numbers x whose representations begin with the same $M_r - 1$ digits determine numbers $f(x)$ whose representations begin with the same n_r digits, and hence, for such numbers $f(x)$,

$$0 \leqslant f(x) - f(\xi) < 10^{-n_r}.$$

Since the *greatest* n_r has been chosen, $n_r \to \infty$ as $M_r \to \infty$, and so $f(x) \to f(\xi)$ as $x \to \xi+$. On the other side of ξ, by beginning with an infinite set of digits not all *zero*, we prove similarly that $f(x) \to f(\xi)$ as $x \to \xi-$. There is therefore continuity at $x = \xi$.

On the other hand, suppose that ξ has the two representations $0{\cdot}a_1 a_2 a_3 \ldots$ and $0{\cdot}a_1' a_2' a_3' \ldots$. These give rise to possible representations $0{\cdot}b_1 b_2 b_3 \ldots$ and $0{\cdot}b_1' b_2' b_3' \ldots$ of $f(\xi)$. By the same argument these are the two limits $f(\xi+), f(\xi-)$ as x approaches ξ from either side. It is not easy to devise a scheme of definition $b_n = \phi(a_1, a_2, \ldots, a_m)$ such that every double representation of x leads to two decimals $0{\cdot}b_1 b_2 b_3 \ldots$ and $0{\cdot}b_1' b_2' b_3' \ldots$ representing the same number. There will thus, in general, be a simple discontinuity at all numbers x represented by terminating decimals. Consider for instance the cases

(i) $b_n = a_n + (-1)^{a_n}$,

(ii) $b_n = \min(a_1, \ldots, a_n)$,

(iii) $b_n = 9 - a_n$.

† Cf. Burchnall and Chaundy, *Proc. London Math. Soc.* (2) 24 (1924), 150–7.

EXAMPLES II

1. Discuss the convergence at $x = 0$ of the functions

$$|x|^x, \qquad |x|^{|x|}, \qquad |x|^{\sin(1/x)}, \qquad |\sin(1/x)|^x.$$

2. Show that two continuous functions which have the same values at all rational points are identical.

3. If $f_1(x),\ldots, f_n(x)$ are all continuous at $x = a$ and if, for every ξ, $F(\xi)$ denotes the algebraically greatest of the numbers $f_1(\xi),\ldots, f_n(\xi)$, prove that $F(x)$ is continuous at $x = a$.

4. Discuss the continuity of the functions

$$\begin{aligned} &\text{(i) } f(x) = a && (x \text{ irrational}), \\ &\qquad\quad\;\; = b \neq a && (x \text{ rational}); \\ &\text{(ii) } f(x) = 0 && (x \text{ irrational}), \\ &\qquad\quad\;\; = 1/q && (x \text{ rational and in lowest terms } \pm p/q). \end{aligned}$$

5. Every number lying in $(0, 1)$ is expressed as a radix fraction in scale 2 in the form

$$x = 0 \cdot a_1 a_2 a_3 \ldots.$$

The corresponding value of $f(x)$ is defined as the radix fraction in scale 2

$$f(x) = 0 \cdot b_1 b_2 b_3 \ldots,$$

where
$$b_n = |a_1 - a_n|.$$

Show that $f(x)$ is everywhere continuous.

6. Any number lying in $(0, 1)$ is written as the proper continued fraction

$$x = \frac{1}{a_1+}\,\frac{1}{a_2+}\cdots\frac{1}{a_n+}\cdots.$$

The corresponding value of $f(x)$ is defined as the continued fraction

$$f(x) = \frac{1}{b_1+}\,\frac{1}{b_2+}\cdots\frac{1}{b_n+}\cdots,$$

where $b_n = \phi_n(a_1, a_2,\ldots, a_n)$.

Show, that in general, $f(x)$ is discontinuous at every rational point in $(0, 1)$. Consider in particular the cases

$$\begin{aligned} &\text{(i) } b_n = 1 + a_n, \\ &\text{(ii) } b_n = a_1 + a_2 + \cdots + a_n, \\ &\text{(iii) } b_n = a_{n-1}, \quad b_1 = 1. \end{aligned}$$

7. Discuss the continuity of the function

$$f(x) = q - px,$$

where p/q is the nth convergent to x expressed as a continued fraction, or the last convergent if the continued fraction has less than n convergents, n being a number independent of x.

8. Show that the condition

$$|f(\xi + h) - f(\xi - h)| \to 0 \quad \text{as} \quad h \to 0,$$

although necessary, is not sufficient for the convergence of $f(x)$ at $x = \xi$.

Show more generally that the condition

$$|f(\xi + ph) - f(\xi + qh)| \to 0 \quad \text{as} \quad h \to 0$$

is not sufficient for the convergence of $f(x)$ at $x = \xi$, where p, q are any fixed numbers of either sign.

9. Discuss the existence and continuity at $x = \pm 1$ of the limit-functions of the sequences

$$x^n, \qquad x^n(1-x^2), \qquad \frac{x^{2n+1}+1}{x^{2n}+1}.$$

Discuss also the uniformity of convergence.

10. A sequence $s(n, x)$ is defined by the formula

$$s(n,x) = 0 \qquad (x \text{ irrational}),$$
$$s(n,x) = \sin(n!\,x\pi) \qquad (x \text{ rational}).$$

Show that $s(n, x)$ is discontinuous in x save at points isolated for given n, but that the sequence converges to a sum function everywhere continuous.

11. The series

$$\sum^{\infty} \frac{1}{n^2}\sin\frac{n}{nx-1}$$

is uniformly convergent, but there is no neighbourhood of $x = 0$ throughout which every term of the series is continuous. Is the sum-function continuous at $x = 0$?

[At $x = 1/n$, the nth term, which becomes meaningless, is to be replaced by zero.]

12. If $s(n, x) \to s(x)$ where every $s(n, x)$ and $s(x)$ are continuous at $x = \xi$, show that the continuity of $s(n, x)$ may be said to be 'vanishingly uniform in n' in the sense that

$$|s(n,x)-s(n,\xi)| < \text{a given } \epsilon$$

for every† x such that $\qquad |x-\xi| < \text{some } \delta(\epsilon, \xi)$

and every $\qquad n > \text{some } N(\epsilon, \xi, x),$

the order of choice being ϵ, δ, n.

Prove conversely that, if the continuity of $s(n, x)$ at $x = \xi$ is vanishingly uniform in n in this sense, then $s(x)$ is continuous at $x = \xi$.

13. If $S(x) = \sin x + \frac{1}{2}\sin 2x + \frac{1}{3}\sin 3x + \dots$ to ∞ and p, q are any two integers (zero excluded), prove that $S(px) - S(qx)$ is everywhere continuous.

14. If $S(x) \equiv a_1 \sin x + a_2 \sin 2x + a_3 \sin 3x + \dots$ to ∞, and na_n steadily decreases to a positive limit a, prove that $S(x) \to \frac{1}{2}\pi a$, as $x \to 0$.

15. With the notation of example 14, if $a_n \to 0$ monotonically, prove that $S(x) - S(px)$ is continuous at $x = 0$, where p is any fixed integer (zero excluded).

16. Prove, without using the formula for the sum function of the infinite series

$$\sin x + \tfrac{1}{2}\sin 2x + \tfrac{1}{3}\sin 3x + \dots,$$

that this sum function is discontinuous at $x = 0$.

17. If $\qquad s_n \equiv a_0 + a_1 + \dots + a_n \to s,$

and if $\qquad s(x) = a_0 + a_1 x + \dots + a_n x^n + \dots$ to ∞,

show that

$$S(n,x) \equiv (s-s_0)+(s-s_1)x+\dots+(s-s_n)x^n \to \frac{s-s(x)}{1-x} \quad \text{as } n \to \infty,$$

if $-1 < x < 1$.

If further, as $n \to \infty$, $S(n, 1)$ converges (to S say), show that

$$\lim_{x\to 1} \frac{s(x)-s}{x-1}$$

exists and equals S.

† Strict analogy with the vanishingly uniform convergence of § 8 would require 'some' to replace 'every' here. The condition is then only sufficient for convergence of $s(x)$ to $s(\xi)$ *along some sequence of approach* (composed of these points x).

18. With the notation of example 17, show that, if $a_n = (-1)^{n-1/n}$, then $s(n, 1)$ converges to the sum $\frac{1}{2}$, the sequence $|s-s_n|$ being monotonic.

19. Prove that for a certain range of values of x the series

$$\log 2 \cos(\tfrac{1}{2}x+\alpha)+(\log 2-1)\cos(\tfrac{3}{2}x+\alpha)+(\log 2-1+\tfrac{1}{2})\cos(\tfrac{5}{2}x+\alpha)+\ldots$$

has the sum
$$\frac{x\cos\alpha-\sin\alpha\log\sec^2\frac{1}{2}x}{4\sin\frac{1}{2}x}.$$

Discuss the continuity and uniform convergence.

20. If $$s(n,x) = a_0+a_1x+\ldots+a_nx^n \to s(x) \text{ in } (0, 1($$

and if $$s(x)\to s \quad \text{as} \quad x\to 1,$$

prove that $$s(n, 1)\to s \quad \text{as} \quad n\to\infty,$$

(i) if every a_n is positive,

or (ii) if $na_n \to 0$. [TAUBER.]†

† *Monatshefte f. Math. u. Phys.*, 8 (1897), 273–7, quoted in Knopp, *Theory and Application of Infinite Series* (1928), 501.

III

THE CONTINUOUS FUNCTION (*continued*)

1. Continuous functions as bounded functions

In the preceding chapter we discussed certain properties of functions continuous at a point. We further saw that the elementary functions are continuous wherever they are defined. This brings us to consider the properties of a function continuous throughout an interval. If the interval is closed, such a function has certain important properties, which are most conveniently stated in the language of the theory of infinite sets.

In that theory a set of numbers $\{x\}$ is said to be 'bounded above', if every number of the set is *algebraically* less than some number A, fixed for the set, i.e. if the inequality

$$x < A$$

is satisfied by every number of the set. Such a number A is called a 'superior bound' of the set. Evidently any number greater than A is also a superior bound, and we can, in fact, consider the set of superior bounds $\{A\}$.

If the set $\{x\}$ has a greatest number x_0, the set is certainly bounded above. The set, however, may be bounded above and yet have no greatest number. In that case it can be shown that the set of superior bounds $\{A\}$ has a least number A_0.

This least superior bound A_0 is called the 'upper bound' of the set $\{x\}$. It is such that

$$\text{every } x < A_0,$$
$$\text{not every } x < A_0-\epsilon,$$

where ϵ is a positive number no matter how small, i.e. given any such ϵ, we can find an $x(\epsilon)$ of the set exceeding $A_0-\epsilon$.

'Inferior bounds', 'lower bound', and 'bounded below' are similarly defined. A set which is bounded both above and below is said simply to be 'bounded'. Thus a bounded set satisfies a double inequality

$$B < x < A.$$

If M is the greater of $|A|$, $|B|$, we can write this double inequality more concisely as

$$|x| < M.$$

We fit this theory to our present purpose by considering the set of values $\{f(x)\}$ assumed by a function $f(x)$ in some domain of definition.

If the set is bounded, or bounded above (below), we speak of the function as bounded, or bounded above (below) in the domain.

If $f(x)$ is continuous at ξ, the function is bounded in the neighbourhood of ξ, for, by the definition of continuity,

$$f(\xi)-\epsilon < f(x) < f(\xi)+\epsilon$$

throughout some interval

$$|x-\xi| < \delta(\epsilon, \xi)$$

which surrounds the point ξ. More generally

(1) *$f(x)$ is bounded in any closed interval throughout which it is continuous.*

For let $f(x)$ be continuous throughout (a, b), where, for clearness, we may suppose $a < b$. Then, as we have just seen, $f(x)$ being continuous at a is certainly bounded in some interval $(a, a+\delta)$. Here δ is not defined precisely, but only by what is tantamount to an inequality, for we may obviously replace δ by any smaller positive δ'. We can thus consider the set of numbers $\{\xi\}$ which belong to (a, b) and are such that $f(x)$ is bounded in (a, ξ).

Then evidently every $\xi \leq b$. If b itself is a number of the set, we at once have what we want. If not, the set $\{\xi\}$ is still bounded and so has either a greatest number or else an upper bound ξ_0, such that, for any positive ϵ, ϵ', sufficiently small†

$$f(x) \text{ is bounded in } (a, \xi_0-\epsilon),$$
$$f(x) \text{ is unbounded in } (a, \xi_0+\epsilon').$$

Consequently $f(x)$ must be unbounded in every $(\xi_0-\epsilon, \xi_0+\epsilon')$. But this is impossible, since by hypothesis $f(x)$ is continuous at ξ_0. We are thus driven back to the conclusion that $f(x)$ is bounded in (a, b).

It is important to observe that the continuity must extend over the *closed* interval (a, b). For example, the function $f(x) \equiv x^{-1}$ is continuous over the half-open interval $)0, 1)$, but is unbounded in that interval.

2. The bounding values of a continuous function

A function continuous throughout a closed interval, since it is bounded in the interval, has either a greatest value or an upper bound in the interval and, similarly, either a least value or a lower bound. It is characteristic of a continuous function that it actually reaches its bounding values, i.e. that

(2) *A function continuous throughout a closed interval attains in that interval a greatest value and a least value.*

† In the extreme case in which b itself is the upper bound of $\{\xi\}$ we must take ϵ' zero in order to keep our discussion within the boundaries of the interval (a, b).

For suppose that $f(x)$ is continuous in (a, b) and that it has no greatest value in (a, b) but only an upper bound M, i.e. suppose that

$$\text{every } f(x) < M,$$
$$\text{some } f(x_\epsilon) > \text{any } M-\epsilon.$$

Consider the function

$$F(x) \equiv \{M-f(x)\}^{-1}.$$

Then, by the hypothesis concerning M,

$$\text{some } F(x_\epsilon) > \text{an arbitrary } \epsilon^{-1},$$

and thus $F(x)$ is unbounded in (a, b). But, since $f(x)$ is continuous throughout (a, b), it follows from chapter II § 4 (17) that $F(x)$ is continuous throughout (a, b) except at zeros of $M-f(x)$. Of these by hypothesis there are none, for $f(x) \neq M$ in the interval. Accordingly $F(x)$ is continuous throughout (a, b) and therefore, by (1), is bounded in (a, b). The hypothesis of an upper bound has thus led us to a contradiction.

The function $f(x)$ has therefore a greatest value in (a, b). The function $-f(x)$ has likewise a greatest value in the interval: that is to say, $f(x)$ itself has also a least value in the interval.

The argument, of course, applies only to a closed interval (a, b). For example, the function $f(x) \equiv x$ is continuous in $)0, 1($. It has upper and lower bounds $1, 0$ in that interval, but attains neither of them in the open interval.

A point at which a continuous function attains its greatest (least) value in the interval may be called 'a point of greatest (least) value in the interval'. There may be many such points or even an infinite set of them.

Thus $f(x) \equiv x^2 \sin^2(\pi/x)$ is continuous throughout $(-1, 1)$ and has in that interval the least value zero, which it assumes at all points of the infinite set

$$\pm 1, \pm\frac{1}{2}, \ldots, \pm\frac{1}{n}, \ldots.$$

If ξ_1, ξ_2 be points of greatest and least value respectively in (a, b), then we have

$$f(\xi_2) \leqslant f(x) \leqslant f(\xi_1)$$

throughout (a, b). We must include in this theory of greatest and least values the limiting case in which $f(\xi_1) = f(\xi_2)$, and in which therefore we have

$$f(\xi_2) = f(x) = f(\xi_1),$$

the function being constant throughout the interval.

3. Maxima and minima of continuous functions

Points of greatest (least) value in an interval that are not end-points of the interval are distinguished as 'points of maximum (minimum) value' or briefly as 'maxima (minima)'. More generally we say that ξ is a maximum (minimum) of $f(x)$, if $f(\xi)$ is the greatest (least) value of $f(x)$ in some neighbourhood of ξ, say the interval $(\xi-\delta, \xi+\delta)$. Thus in a given interval there may be several maxima (minima), which are not necessarily all (or any of them) points of greatest (least) value in the interval, for the points of greatest and of least value may be end-points. In like fashion a point of greatest (least) value in an interval may be an end-point of the interval and therefore not necessarily a maximum (minimum).

We may illustrate these points by considering the function

$$f(x) = x\sin x + \cos x$$

in the interval $(-\pi, \pi)$, and in particular its values

$$f(0) = 1, \qquad f(\tfrac{1}{2}\pi) = \tfrac{1}{2}\pi, \qquad f(\pi) = -1.$$

Now, by the fundamental trigonometric inequality, $x \geqslant \sin x$ in $(0, \frac{1}{2}\pi)$, and thus, in $(0, \frac{1}{2}\pi)$,

$$x\sin x + \cos x \geqslant \sin^2 x + \cos x \geqslant \sin^2 x + \cos^2 x,$$

i.e. $$f(x) \geqslant f(0) \quad \text{in} \quad (0, \tfrac{1}{2}\pi),$$

and so too in $(0, -\frac{1}{2}\pi)$, since $f(-x) = f(x)$. That is to say, $f(0)$ is a least value in the interval $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$ and consequently is a minimum value of $f(x)$, but it is not a least value in the interval $(-\pi, \pi)$, since $f(0) > f(\pi)$.

Again

$$f(x)-f(\pi) = x\sin x+(1+\cos x) \geqslant 0 \quad \text{in} \quad (-\pi, \pi),$$

for in this interval $x\sin x$, $1+\cos x$ are neither negative. Thus $f(\pi)$ and similarly $f(-\pi)$ are least values in the interval $(-\pi, \pi)$. But they are not minima. For

$$f(\pi)-f(x+\pi) = (x+\pi)\sin x+\cos x-1 > 0 \quad \text{in} \quad (0, \tfrac{1}{2}\pi),$$

since we have already shown that

$$x\sin x + \cos x-1 \geqslant 0 \quad \text{in} \quad (0, \tfrac{1}{2}\pi).$$

Thus $$f(\pi) \geqslant f(x) \quad \text{in} \quad (\pi, \tfrac{3}{2}\pi),$$

and so $f(\pi)$ is not a least value of $f(x)$ in any neighbourhood of $x = \pi$: in other words, is not a minimum of $f(x)$. Finally

$$f(\tfrac{1}{2}\pi)-f(x+\tfrac{1}{2}\pi) = \tfrac{1}{2}\pi(1-\cos x)+\cos x(\tan x-x) > 0 \quad \text{in} \quad (0, \tfrac{1}{2}\pi),$$

since $\tan x-x \geqslant 0$ in $(0, \frac{1}{2}\pi)$ by the fundamental trigonometric inequality. Similarly

$$f(\tfrac{1}{2}\pi)-f(\tfrac{1}{2}\pi-x) = (\tfrac{1}{2}\pi-x)(1-\cos x)+x-\sin x \geqslant 0 \quad \text{in} \quad (0, \tfrac{1}{2}\pi),$$

again by the fundamental trigonometric inequality. Thus $f(\frac{1}{2}\pi)$ is the greatest value of $f(x)$ in $(0, \pi)$ and is therefore a maximum, since $(0, \pi)$ is a neighbourhood of $\frac{1}{2}\pi$.

We have thus found for the function $x\sin x+\cos x$ a minimum which is *not* a least value, a least value which is not a minimum, and a greatest value which is also a maximum. These facts are, of course, more easily established by the methods of the next chapter and are more easily appreciated from a graph of the function.

It is important to distinguish maxima (minima) from points of greatest (least) value in an interval. The maxima of a function, from the nature of the definition, are intrinsically connected with the function. But, if an interval be taken in which a function has no maximum, the greatest value occurs at an end-point and therefore depends on the choice of interval and not intrinsically on the function itself.

It follows at once from the definition that

(3) *Between two maxima must occur a minimum.*

Let ξ_1, ξ_2 be two maxima and, say, $\xi_1 < \xi_2$. Then

$$f(\xi_1) > f(x) \text{ in some interval } (\xi_1, \xi_1+\delta_1),$$
$$f(\xi_2) > f(x) \text{ in some interval } (\xi_2-\delta_2, \xi_2)$$

Thus neither $f(\xi_1)$ nor $f(\xi_2)$ can be the least value of $f(x)$ in (ξ_1, ξ_2) The point of least value must be internal and so is a minimum.† Similarly, between two minima must occur a maximum.

With most functions that arise naturally maxima and minima occur at isolated points and we may then say simply that

(4) *Maxima and minima occur alternately.*

4. Monotonic functions

The ideas of maxima and minima are intimately related to that of 'monotonic'. A function is said to be *monotonic* in a domain if

$$\frac{f(x_1)-f(x_2)}{x_1-x_2} \tag{5}$$

has the same sign for every pair of points x_1, x_2 of the domain. If this sign is positive, i.e. if

$$f(x_1) < f(x_2), \quad \text{whenever} \quad x_1 < x_2, \tag{6}$$

we say that $f(x)$ is 'ascending monotonic' or 'steadily increasing' in the domain. So, if the constant sign is negative, i.e. if

$$f(x_1) > f(x_2), \quad \text{whenever} \quad x_1 < x_2, \tag{7}$$

† Even in the extreme case in which $f(\xi_1) = f(x) = f(\xi_2)$, every point in (ξ_1, ξ_2) is a point of greatest and also of least value. Every point in $)\xi_1, \xi_2($ therefore counts as a minimum (as well as a maximum) and the argument is not invalidated

we say that $f(x)$ is 'descending monotonic' or 'steadily decreasing' in the domain.

We must observe that these inequalities (6), (7) are to be taken strictly and that they rule out the possibility $f(x_1) = f(x_2)$ in the domain.

If this possibility be admitted, we must have

$$f(x_1) = f(x) = f(x_2)$$

for every x between $f(x_1)$ and $f(x_2)$, since otherwise

$$\frac{f(x_1)-f(x)}{x_1-x} \quad \text{and} \quad \frac{f(x_2)-f(x)}{x_2-x}$$

have opposite signs. The function must therefore be constant over this part of the domain, and over the whole domain we may call the function 'monotonous'.

Thus $f(x)$ is 'ascending monotonous' or 'never decreasing', if the ratio (5) is never negative, and 'descending monotonous' or 'never increasing' if the ratio (5) is never positive.

We connect 'monotonic' with 'maxima' and 'minima' by the theorem:

(8) *A function which is continuous in (a, b) and has neither maximum nor minimum in $)a, b($ is monotonic in (a, b).*

For the points of greatest and of least value in (a, b) must be at the end-points. It is unimportant which is which, and we may suppose that

$$f(a) \leqq f(x) \leqq f(b)$$

throughout the interval. In the extreme case

$$f(a) = f(x) = f(b) = c, \text{ say},$$

in which the function is constant throughout the interval, it assumes its greatest value c (as well as its least value c) at every point of the interval. Every point of $)a, b($ would therefore strictly be both a maximum and a minimum, a possibility excluded by hypothesis.†

We have then $\quad f(a) < f(x_1) < f(b)$

when $a < x_1 < b$. Hence

$$\frac{f(x_1)-f(b)}{x_1-b} \quad \text{has the sign of} \quad \frac{f(a)-f(b)}{a-b}.$$

By applying the same argument to the sub-interval (x_1, b), we have

$$f(x_1) < f(x_2) < b$$

when $x_1 < x_2 < b$, and thus

$$\frac{f(x_1)-f(x_2)}{x_1-x_2} \quad \text{has the sign of} \quad \frac{f(x_1)-f(b)}{x_1-b}$$

† If this strict interpretation of maximum and minimum is not desired we must read 'monotonous' for 'monotonic' in (8).

and so of $$\frac{f(a)-f(b)}{a-b}.$$

Hence $\{f(x_1)-f(x_2)\}/(x_1-x_2)$ has the same sign throughout (a,b) and the function is therefore monotonic in the interval.

Thus the function, if it has isolated maxima and minima, is monotonic between these points, ascending steadily from minimum to maximum and descending steadily from maximum to minimum.

This theory of maxima and minima is fundamental in the Differential Calculus and we return to it in subsequent chapters.

5. Inverse functions of continuous functions

The theory also bears closely on the theory of inverse functions, as will appear from the following theorem:

(9) *If in an interval a function is one-valued, continuous, and ascending monotonic, its inverse is likewise one-valued, continuous, and ascending monotonic in the corresponding interval.*

Suppose that $y = y(x)$ is one-valued, continuous, and ascending monotonic in (a,b), where to fix ideas we take $a < b$.

Then, for every pair of points x_1, x_2, the ratio
$$\frac{y_1-y_2}{x_1-x_2}$$
is positive (and never zero). Hence we cannot have $y_1 = y_2$, unless $x_1 = x_2$. The inverse function $x = x(y)$ is therefore one-valued. It is also ascending monotonic, since
$$\frac{x_1-x_2}{y_1-y_2}$$
is always positive for every pair of points y_1, y_2 in the interval $\{y(a), y(b)\}$.

For the continuity of the inverse function we need to prove that
$$|x-\xi| < \epsilon \quad \text{whenever} \quad |y-\eta| < \text{some } \delta(\epsilon,\xi),$$
where ξ is any point in (a,b) and $\eta = y(\xi)$. Remove the open interval $|x-\xi| < \epsilon$ from the interval (a,b). There remain the two closed intervals $(a, \xi-\epsilon)$, $(\xi+\epsilon, b)$ in which $y(x)$ is still continuous and therefore also $|y(x)-\eta|$ is continuous and has a least value. This cannot be zero, since $y(x)-\eta$ vanishes only at $x = \xi$. The least value is then some positive δ and so
$$|y-\eta| \geqslant \delta$$
throughout $(a, \xi-\epsilon)$ and $(\xi+\epsilon, b)$.

Hence the condition $|y-\eta| < \delta$ confines x to the interval $|x-\xi| < \epsilon$, which establishes the continuity of $x(y)$ at $y = \eta$.

6. Branches and branch-points

Now a one-valued, continuous function will not, in general, have a one-valued inverse. Thus the function $y = x^2$ has the two-valued inverse $x = \pm\sqrt{y}$, while the function $y = \sin x$ has the many-valued inverse $x = n\pi + (-)^n \sin^{-1} y$. But, if the function have isolated maxima and minima, it can be cut at these points into monotonic segments and, by (9), each of these segments will have a one-valued, continuous, monotonic inverse. The complete inverse function will thus consist of a number of one-valued, continuous segments or *branches* as they are technically called. A many-valued inverse function which can be split up in this way into one-valued, continuous branches is still admitted to the title of continuous function.

Thus the two-valued inverse $x = \pm\sqrt{y}$ can be split up into the two continuous branches

$$x = +|y^{\frac{1}{2}}| \quad \text{(the positive branch)}$$

and

$$x = -|y^{\frac{1}{2}}| \quad \text{(the negative branch)},$$

the point of separation being $x = 0$, $y = 0$, which is the solitary minimum of $y = x^2$.

The many-valued inverse $x = n\pi + (-)^n \sin^{-1} y$ can be split up into the continuous branches

$$x = \sin^{-1} y, \qquad \begin{aligned} &x = \pi - \sin^{-1} y, \quad &x = 2\pi - \sin^{-1} y, \text{ etc.},\\ &x = -\pi - \sin^{-1} y, \quad &x = -2\pi - \sin^{-1} y, \text{ etc.},\end{aligned}$$

where $\sin^{-1}$ denotes, as usual, the principal value, i.e. the appropriate positive or negative acute angle. The points of separation are at $x = (n+\frac{1}{2})\pi$, $y = \pm 1$, the maxima and minima of $y = \sin x$.

The maxima and minima of the original function $y = y(x)$ are thus points at which branches of the inverse function $x = x(y)$ are connected up. They are known as the 'branch-points' of the inverse function. In other words, 'branch-points' of a many-valued function are points at which two or more branches have the same value. Thus $y = 0$ is the solitary branch-point of the two-valued function $x = \sqrt{y}$, for there the two branches $x = \pm|y^{\frac{1}{2}}|$ have the same value zero.

The several branches of the inverse function need not, of course, have the same domain of definition. Thus

$$y(x) \equiv I(x)\sin^2 \pi x$$

defines a continuous function which is constant over the interval (0, 1).

Elsewhere the inverse function exists and, as is readily seen from a figure, has infinitely many branches of the type

$$x = n+\frac{1}{\pi}\sin^{-1}\sqrt{\frac{y}{n}}, \qquad x = n+1-\frac{1}{\pi}\sin^{-1}\sqrt{\frac{y}{n}},$$

where n is an integer.

The domain of each of the typical branches written above is the closed interval $0 \leqslant y \leqslant n$. The branches as a whole have therefore only the point $y = 0$ common to their domains of definition. Other examples could be devised in which even this common element was lacking. It thus appears that a k-valued inverse $x = x(y)$ need never, in the field of the real variable, have as many as k values of x for any value of y. The number k, in fact, measures the multiplicity of *branches* and not the multiplicity of *values*.

Inverse functions are merely special cases of functions defined implicitly, and the theory of continuous branches extends without difficulty to functions so defined. Consider as a simple example the implicit relation

$$x^2+y^2 = 1.$$

The extreme values of x and y alike are ± 1. The function $y = y(x)$ defined implicitly by this relation is composed of the two continuous branches

$$y = (1-x^2)^{\frac{1}{2}} \quad \text{and} \quad y = -(1-x^2)^{\frac{1}{2}}$$

defined over the common domain $(-1, 1)$, the end-points $x = \pm 1$ being branch-points. The implicit function $x = x(y)$ is similarly composed of the two continuous branches

$$x = (1-y^2)^{\frac{1}{2}} \quad \text{and} \quad x = -(1-y^2)^{\frac{1}{2}}.$$

The theory of implicit functions is more fully considered in a subsequent chapter (XI).

It should be added that the foregoing theory can be taken over to many-valued functions in general; in actual practice many-valued functions occur most naturally as implicit functions.

7. The continuous function as an unbroken function

The foregoing theory of maxima and minima, monotonic and inverse functions can be studied very easily with the aid of a figure. The theory, in fact, largely consists of ready inductions from the typical graph of a continuous function.

There is another fundamental induction from such a graph to which we now come. It is that a continuous curve in going from one value of y to another must pass through every intermediate value of y: in other words, that the curve is all of one piece. We may say that such

a curve is 'unbroken': in earlier days mathematicians would simply have said that it was 'continuous', for this graphical continuity represents an earlier notion of continuity. The idea is, however, less fundamental analytically and less comprehensive: as we shall see later, a curve may be unbroken and yet not be continuous in the modern sense. We now prove conversely that a continuous curve $y = f(x)$ is unbroken. We prove it in the rather more special form:

(10) *If a function is continuous throughout a closed interval and has opposite signs at the end-points, then it must vanish at some point of the interval.*

Let $f(x)$ be continuous in (a, b) and, to fix ideas, suppose that $f(a) > 0$, $f(b) < 0$. If $f(x)$ do not vanish at a point ξ of the interval, then, in virtue of the continuity at that point, we have, by taking $\frac{1}{2}|f(\xi)|$ as the ϵ of the definition, that

$$f(\xi)-\tfrac{1}{2}|f(\xi)| < f(x) < f(\xi)+\tfrac{1}{2}|f(\xi)|$$

throughout some interval $|x-\xi| < \delta(\xi)$, i.e. $f(x)$ keeps the sign of $f(\xi)$ throughout some neighbourhood of ξ. Hence, if $f(x)$ changes sign in every neighbourhood of a point ξ_0, we must have $f(\xi_0) = 0$.

In particular then, since, by hypothesis, $f(a)$ is positive, $f(x)$ is also positive throughout some interval $(a, a+\delta)$. But, by hypothesis, $f(x)$ is not positive throughout the complete interval (a, b). We can therefore consider the set of points ξ of (a, b) such that $f(x)$ is positive throughout (a, ξ). The set is clearly bounded above (e.g. by the number b) and so has either a greatest number ξ_0 or else an upper bound ξ_0. In either case we can associate with it a number ξ_0 such that however small ϵ, ϵ'

$$f(x) \text{ is positive throughout } (a, \xi_0-\epsilon),$$

$$f(x) \text{ changes sign in } (a, \xi_0+\epsilon').$$

Hence $f(x)$ changes sign in every $(\xi_0-\epsilon, \xi_0+\epsilon')$, i.e. in every neighbourhood of ξ_0. Thus, by what we have proved above, $f(\xi_0) = 0$.

It should be noticed that we have proved not only that ξ_0 is a zero of $f(x)$, but also that it is the *first* zero reckoning from a. We could similarly determine the *last* zero reckoning from a, i.e. the first reckoning from b.

We may state the theorem in less precise form:

(11) *A continuous function can change sign only by passing through zero.*

So stated it is recognizable as a characteristic property of the continuous phenomena of our experience.

From (10) we at once deduce the theorem in its more general form:

(12) *A function continuous in a closed interval assumes therein every value which lies between its values at the end-points of the interval.*

For, if $f(x)$ be continuous throughout (a,b), and if h be any number between $f(a)$ and $f(b)$, then $f(x)-h$ is continuous in (a,b) and has opposite signs at $x = a$ and $x = b$. Thus, by (10), it vanishes somewhere in the interval, i.e., somewhere in the interval, $f(x) = h$.

More generally, if ξ_1, ξ_2 are respectively points of greatest value and least value in the interval, we can apply (12) to the sub-interval (ξ_1, ξ_2) and so deduce:

(13) *A function continuous in a closed interval assumes therein every value lying between its greatest and least values in the interval.*

The effect of (12) is that y runs through all values in an interval, if x runs through all values in an interval in which $y(x)$ is continuous. In other words, a continuous function $y(x)$ associates *continuous variation* in y with *continuous variation* in x.

8. Uniform continuity

If $f(x)$ is continuous throughout a domain, the speed with which $f(x)$ converges to $f(\xi)$ will vary from point to point of the domain. In chapter I § 10, we have considered the varying speed with which $s(n,x)$, a sequence of variable terms, converges to its limit-function $s(x)$. There is strict analogy between these two cases, but it is somewhat masked, because, in the sequence, the speed of convergence varies on account of the intrusion of a new variable x distinct from the parameter of convergence n. In the function, the disturbing variable is the parameter of convergence itself.

We perhaps make the analogy more visible, if, in the function, we write $x = \xi+t$, and consider always the continuity of $f(\xi+t)$ at $t = 0$, separating the parameter of convergence t from ξ the variable on which depends the speed of convergence.

In the sequence we estimate the speed of convergence by noting $N(\epsilon, x)$ the ordinal number of the term $s(N,x)$ after which consistently

$$|s(n,x)-s(x)| < \epsilon.$$

So, in the function, we measure the closeness of the continuity by the breadth of the interval $|x-\xi| < \delta(\epsilon,\xi)$ throughout which

$$|f(x)-f(\xi)| < \epsilon.$$

Pursuing the analogy, we say that $f(x)$ is *uniformly continuous* over a domain, if, given ϵ, we can find $\delta(\epsilon)$ *independent of* ξ such that

$$|x-\xi| < \delta \quad \text{secures} \quad |f(x)-f(\xi)| < \epsilon,$$

that is to say, if $f(x_1)$, $f(x_2)$ differ by less than ϵ whenever x_1, x_2 differ by less than some corresponding $\delta(\epsilon)$.

The idea of the uniform continuity of a function introduces few difficulties as compared with those that arise from the uniform convergence of sequences, because, in general terms, continuity implies uniform continuity. Precisely

(14) *A function continuous throughout a closed interval is uniformly continuous over it.*

If $f(x)$ is continuous throughout (a,b) and ξ is a point of the interval, we have $\quad |f(x_1)-f(x_2)| \leqslant \epsilon, \quad \text{if} \quad |x-\xi| \leqslant \text{some } \delta(\epsilon,\xi).$

To make δ precise for given ϵ, ξ we take its greatest value or upper bound δ_0 and so define the one-valued function $\delta_0(\epsilon, x)$, which, for given ϵ, we shall show to be a continuous function of x.

For take $\theta(\epsilon,\xi)$ such that $|\theta| < \delta_0(\epsilon,\xi)$. Then the interval

$$|x-(\xi+\theta)| \leqslant \delta_0-|\theta| \tag{15}$$

forms part of the interval

$$|x-\xi| \leqslant \delta_0,$$

and so $|f(x_1)-f(x_2)| \leqslant \epsilon$ throughout (15). Thus $\delta_0-|\theta|$ is a possible δ at $\xi+\theta$, i.e.

$$\delta_0(\epsilon,\xi)-|\theta| \leqslant \delta_0(\epsilon,\xi+\theta), \quad \text{if} \quad |\theta| \leqslant \delta_0(\epsilon,\xi).$$

So, by proceeding from $\xi+\theta$ to ξ, we have

$$\delta_0(\epsilon,\xi+\theta)-|\theta| \leqslant \delta_0(\epsilon,\xi), \quad \text{if} \quad |\theta| \leqslant \delta_0(\epsilon,\xi+\theta).$$

We secure this inequality for θ, if $|\theta| \leqslant \frac{1}{2}\delta_0(\epsilon,\xi)$, and hence

$$|\delta_0(\epsilon,\xi+\theta)-\delta_0(\epsilon,\xi)| \leqslant |\theta|, \quad \text{if} \quad |\theta| \leqslant \tfrac{1}{2}\delta_0(\epsilon,\xi),$$

i.e. $\quad |\delta_0(\epsilon,x)-\delta_0(\epsilon,\xi)| \leqslant |x-\xi|, \quad \text{if} \quad |x-\xi| \leqslant \frac{1}{2}\delta_0(\epsilon,\xi).$

This establishes the continuity of $\delta_0(\epsilon,x)$, regarded as a function of x, throughout (a,b). It has therefore a least value $\delta_0'(\epsilon)$ in the interval, which is positive (and not zero), since it is the value of $\delta_0(\epsilon,\xi)$ at some ξ. Hence

$$|f(x_1)-f(x_2)| \leqslant \epsilon,$$

if x_1, x_2 lie in the same interval $|x-\xi| \leqslant \delta_0'(\epsilon)$, i.e. if

$$|x_1-x_2| \leqslant 2\delta_0'(\epsilon).$$

This proves the continuity uniform over (a,b).

We should observe that the continuity must be known to extend

throughout the *closed* interval, or else the theorem is false. Thus $f(x) \equiv x^{-1}$ is continuous throughout $)0, 1)$, but is not uniformly continuous over this half-open interval.

9. Convergence of functions of many variables

The theory of the convergence and continuity of functions of many variables is sufficiently illustrated by considering $f(x, y)$ a function of two variables. We say that

$$f(x,y) \text{ converges at } (\xi, \eta) \text{ to the limit } A,$$

if $|f(x,y)-A| <$ any positive ϵ throughout some associated interval

$$|x-\xi| < \delta(\epsilon, \xi, \eta), \qquad |y-\eta| < \delta'(\epsilon, \xi, \eta).$$

Suppose now that this condition is satisfied and that, in addition, the functions $x(t)$, $y(t)$ converge respectively to ξ, η as t converges to some τ; then it follows that $f\{x(t), y(t)\}$ converges to A, as t converges to τ. For

$$|f\{x(t), y(t)\}-A| < \epsilon$$

is secured by

$$|x(t)-\xi| < \delta(\epsilon, \xi, \eta), \qquad |y(t)-\eta| < \delta'(\epsilon, \xi, \eta),$$

which in turn are secured by

$$|t-\tau| < \text{some } \theta(\epsilon, \tau).$$

In similar fashion, if the sequences $\{x_n\}, \{y_n\}$ converge respectively to ξ, η, then it follows that the sequence $\{f(x_n, y_n)\}$ converges to A. For again

$$|f(x_n, y_n)-A| < \epsilon$$

is secured by

$$|x_n-\xi| < \delta(\epsilon, \xi, \eta), \qquad |y_n-\eta| < \delta'(\epsilon, \xi, \eta),$$

which in turn are secured by

$$n > \text{some } N(\epsilon).$$

In geometrical language we may speak of $x = x(t)$, $y = y(t)$ as a 'curve' through (ξ, η). Less generally, but more usefully, we consider a 'curve' not merely 'convergent' at (ξ, η) but continuous for every t in some domain. We may call such a curve a 'continuous curve' through (ξ, η). The double sequence $\{x_n\}$, $\{y_n\}$ converging respectively to ξ, η I shall speak of as a 'sequence of approach' to (ξ, η). In these terms we can now state the above pair of results as a generalization of theorem (3) of chapter II:

(16) *If $f(x,y)$ converges to A at (ξ, η), it also converges to A* (i) *along every continuous curve through* (ξ, η), (ii) *along every sequence of approach to* (ξ, η).

The converse also is true, namely that

(17) *$f(x,y)$ converges to A at (ξ,η), if it converges to A either* (i) *along every sequence of approach to (ξ,η) or* (ii) *along every continuous curve through (ξ,η).*

We establish (i) by showing, exactly as in chapter II (4), that, if $f(x,y)$ do not converge to A at (ξ,η), then we can construct $\{x_n, y_n\}$, a sequence of approach to (ξ,η), such that the sequence $\{f(x_n,y_n)\}$ does not converge to A. I suppress the proof.

We prove (ii) by showing that through any such sequence of approach $\{x_n, y_n\}$ and through (ξ,η) itself we can construct a continuous curve $x = x(t)$, $y = y(t)$.

To do this, take any monotonic sequence $\{\tau_n\}$ ascending (say) to τ as its limit, and for the curve take the infinite broken polygon defined by

$$x(t) = \frac{x_n(\tau_{n+1}-t)+x_{n+1}(t-\tau_n)}{\tau_{n+1}-\tau_n}, \qquad y(t) = \frac{y_n(\tau_{n+1}-t)+y_{n+1}(t-\tau_n)}{\tau_{n+1}-\tau_n}$$
$$(\tau_n \leqslant t \leqslant \tau_{n+1})$$

for $n = 1, 2, 3, \ldots$, and finally

$$x(\tau) = \xi, \qquad y(\tau) = \eta.$$

The functions $x(t)$, $y(t)$ are seen to be continuous throughout the closed interval (τ_1, τ).

As in chapter II (6) we can establish the 'general' condition of convergence, i.e. the condition of convergence to an unspecified limit:

(18) *$f(x,y)$ converges at (ξ,η), if corresponding to any ϵ we can find $\delta(\epsilon,\xi,\eta)$, $\delta'(\epsilon,\xi,\eta)$ such that*

$$|f(x_1,y_1)-f(x_2,y_2)| < \epsilon$$

for every pair of points (x_1,y_1), (x_2,y_2) in the interval $|x-\xi| < \delta$, $|y-\eta| < \delta'$, the point (ξ,η) itself excepted.

If the point (ξ,η) be not excepted, this condition becomes the 'general' condition of *continuity* and does not, in effect, differ from the 'special' condition of continuity in which the fundamental equality is

$$|f(x,y)-f(\xi,\eta)| < \epsilon.$$

In (17), we should emphasize that $f(x,y)$ must converge along every *curve* and not merely in every *direction*, i.e. along every straight line through (ξ,η). An example confirms this. Write

$$f(x,y) \equiv \frac{x^2y}{x^4+y^2}. \tag{19}$$

Then, along any straight line $x = r\cos\alpha$, $y = r\sin\alpha$, $f(x,y) \to 0$ as $r \to 0$, but, on the parabola $y = kx^2$, $f(x,y)$ has the constant value $k(1+k^2)^{-1}$. Thus, as we travel to the origin along different parabolas

of this system, $f(x,y)$ converges to distinct limits, and therefore, by (16), is not convergent in the full sense.

Still less then is it sufficient for the convergence of $f(x,y)$ at (ξ,η) to know that
$$\lim_{x\to\xi} f(x,\eta) \quad \text{and} \quad \lim_{y\to\eta} f(\xi,y) \tag{20}$$
exist and are the same, i.e. to know that $f(x,y)$ converges, and to the same limit, merely along the two parallels to the axes through (ξ,η).

It may happen that $f(x,y)$ is not defined along the lines $x=\xi$ and $y=\eta$. In that case we may consider, instead of the two limits (20), the *repeated* limits
$$\lim_{x\to\xi}\lim_{y\to\eta} f(x,y), \qquad \lim_{y\to\eta}\lim_{x\to\xi} f(x,y). \tag{21}$$

If $f(x,y)$ converges in the full sense at (ξ,η), we write its limit as
$$\lim_{x,y\to\xi,\eta} f(x,y), \tag{22}$$
and speak of it as the *double* limit.

By the foregoing argument the existence and equality of the *repeated* limits (21) is insufficient to secure the existence of the *double* limit (22). The example (19), or more simply
$$f(x,y) \equiv \frac{xy}{x^2+y^2} \tag{23}$$
taken at the origin, shows this.

The existence of the double limit, on the other hand, is insufficient to secure the existence of either of the single limits
$$\lim_{y\to\eta} f(x,y) \quad \text{or} \quad \lim_{x\to\xi} f(x,y).$$
This is clear from the example
$$f(x,y) \equiv (x+y)\sin(xy)^{-1} \tag{24}$$
taken at the origin. For
$$|f(x,y)| < |x+y| < \epsilon, \quad \text{if} \quad |x|, |y| < \tfrac{1}{2}\epsilon.$$
But here neither $\lim\limits_{x\to 0}$ nor $\lim\limits_{y\to 0}$ exists.

In this sense the existence of the double limit is insufficient to secure the existence of the repeated limits.

If, however, either of the single limits exists, and if also the double limit exists, then the corresponding repeated limit exists and is equal to the double limit. Precisely,

(25) *If* $\displaystyle\lim_{x,y\to\xi,\eta} f(x,y)$ *exists,*

and if, for every x in the neighbourhood of ξ,
$$\lim_{y\to\eta} f(x,y) \ \textit{exists,}$$
then
$$\lim_{x\to\xi}\lim_{y\to\eta} f(x,y) = \lim_{x,y\to\xi,\eta} f(x,y).$$

For let $\lim_{y \to \eta} f(x,y) = \phi(x)$, so that

$$|f(x,y)-\phi(x)| < \tfrac{1}{2}\epsilon, \quad \text{if} \quad |y-\eta| < \text{some } \delta(\epsilon, x, \eta). \tag{26}$$

But

$$|f(x,y)-A| < \tfrac{1}{2}\epsilon, \quad \text{if} \quad |x-\xi| < \text{some } \delta_1(\epsilon, \xi, \eta)$$
$$\text{and} \quad |y-\eta| < \text{some } \delta_2(\epsilon, \xi, \eta). \tag{27}$$

Thus, if we first choose x such that $|x-\xi| < \delta_1(\epsilon, \xi, \eta)$, and thereupon choose y such that $|y-\eta| < \min\{\delta(\epsilon, x, \eta), \delta_2(\epsilon, \xi, \eta)\}$, then both (26) and (27) hold, and therefore also the inequality

$$|\phi(x)-A| < \epsilon.$$

This is now sufficiently secured by $|x-\xi| < \delta_1$, since y has disappeared, and accordingly

$$\phi(x) \to A \quad \text{as} \quad x \to \xi,$$

and the theorem is proved.

10. Continuity of functions of many variables

The preceding discussion on the convergence of functions of two (or more) variables extends at once to the continuity of such functions, if we take the value at the point as the limit A.

It follows that the continuity of $f(x,y)$ at (ξ, η) is not secured by the continuity at $x = \xi$ of $f(x, \eta)$ regarded as a function of x only, nor by the continuity at $y = \eta$ of $f(\xi, y)$ regarded as a function of y only, nor even by the simultaneous continuity of both these functions of a single variable.

It is convenient to speak of the continuity of $f(x,y)$ (i) 'in (x,y)', (ii) 'in x', (iii) 'in y', according as we are regarding $f(x,y)$ (i) as a function of both variables, (ii) as a function of x only, y being fixed, (iii) as a function of y only, x being fixed.

When we come to consider uniformity of continuity, the continuity of $f(x,y)$ in x (say) may be (i) uniform in x, (ii) uniform in y, (iii) uniform in (x,y), according as, in the inequality

$$|x'-x| < \delta$$

which secures the inequality

$$|f(x',y)-f(x,y)| < \epsilon,$$

δ is (i) independent of x, i.e. is $\delta(\epsilon, y)$;

(ii) independent of y, i.e. is $\delta(\epsilon, x)$;

(iii) independent of both x, y, i.e. is $\delta(\epsilon)$.

By (14), the continuity in x, if it hold over a closed interval, is neces-

sarily uniform in x. But it is not necessarily uniform in y: in fact, in analogy with chapter II (25), we can prove that

(28) *If $f(x,y)$ is continuous in x and in y at (ξ,η), and if the continuity in x (say) is uniform in y near $y=\eta$, then $f(x,y)$ is continuous in (x,y) at (ξ,η).*

For, from the uniform continuity in x,

$$|f(x,y)-f(\xi,y)| < \tfrac{1}{2}\epsilon, \quad \text{if} \quad |x-\xi| < \text{some } \delta_1(\epsilon,\xi,\eta)$$

over some neighbourhood

$$|y-\eta| < \delta_2(\xi,\eta).$$

From the continuity in y

$$|f(\xi,y)-f(\xi,\eta)| < \tfrac{1}{2}\epsilon, \quad \text{if} \quad |y-\eta| < \text{some } \delta_3(\epsilon,\xi,\eta).$$

Hence $$|f(x,y)-f(\xi,\eta)| < \epsilon,$$

if $|x-\xi| < \delta_1$ and $|y-\eta| < \min(\delta_2,\delta_3)$, which establishes the continuity in (x,y).

But we cannot, conversely, infer any such uniformity of continuity merely from the continuity of $f(x,y)$ in (x,y) at a single point. For this continuity gives information only about occurrences within an interval

$$|x-\xi| < \delta_1(\epsilon,\xi,\eta), \qquad |y-\eta| < \delta_2(\epsilon,\xi,\eta),$$

which, in general, becomes vanishingly small as ϵ vanishes. There is thus no *fixed* interval over which we can state the behaviour of $f(x,y)$.

In corroboration consider the example

$$f(x,y) \equiv xy\sum_1^\infty \frac{a_n|y-a_n|}{x^2+(y-a_n)^2} \tag{29}$$

where $\{a_n\}$ is a monotonic, positive sequence such that $\sum\limits_1^\infty a_n$ converges (to s, say).

By direct substitution

$$f(x,0)=0, \qquad f(0,y)=0, \qquad f(0,0)=0.$$

Again, since $$\frac{x|y-a_n|}{x^2+|y-a_n|^2} \leqslant \tfrac{1}{2},$$

the series is comparable for convergence with $\sum a_n$ and so converges everywhere, and, moreover,

$$|f(x,y)| < \tfrac{1}{2}s|y|.$$

Hence $f(x,y)$ is continuous in (x,y) at the origin and is continuous in y at any point $(\xi,0)$.

Again for a fixed y ($\neq a_n$)

$$\sum_1^\infty \frac{ya_n|y-a_n|}{x^2+|y-a_n|^2} < \sum_1^\infty \frac{ya_n}{|y-a_n|},$$

and, since $a_n \to 0$, this is comparable for convergence with $\sum a_n$. It thus converges, say to $\sigma(y)$. Thus, for fixed y ($\neq a_n$), $f(x,y) = x\sigma(y)$; and $f(x, a_n) = 0$. Thus $f(x,y)$ is continuous† in x at any $(0,\eta)$.

Nevertheless this convergence in x is not uniform in y. For uniformity requires that we be able to find some positive h and, given any positive ϵ, some positive $\delta(\epsilon)$ such that $|f(x,y)| < \epsilon$ throughout $|x| < \delta$, $|y| < h$. To show that this is not the case for the function we are considering, choose‡ $a_n < \frac{1}{2}h$ and $\epsilon = \frac{1}{2}a_n^2$. Substitute also $y = a_n+x$ and restrict x to the interval $(0, \frac{1}{2}h)$, so that $|y| < h$. The terms in $f(x,y)$ are all of one sign, and therefore $|f(x,y)|$ exceeds the modulus of any one term. Performing the above substitution in the nth term of the series, we then have

$$|f(x,y)| > \tfrac{1}{2}a_n(a_n+x) > \epsilon,$$

however small x may be. The continuity in x is thus not uniform in y.

There is in actual fact a 'quasi-uniformity' of convergence which lies between the 'vanishingly uniform' convergence of chapter II § 8 and full uniform convergence. The distinction is possibly a little clearer if explained for a sequence $\{s(n,x)\}$ of variable terms. There is *uniform convergence* near $x = \xi$, if throughout some interval $|x-\xi| < \delta$ the fundamental inequality $|s(n,x)-s(x)| < \epsilon$ is secured simply by the inequality $n >$ some $N(\epsilon)$. This represents an order of choice

$$\delta, \quad \epsilon, \quad N.$$

At the other extreme there is *vanishingly uniform convergence* near $x = \xi$, if the fundamental inequality can be secured by the inequalities $n(\epsilon) >$ a prescribed N and $|x-\xi| <$ some $\delta(\epsilon, n)$. This represents an order of choice

$$\epsilon, \quad n, \quad \delta.$$

Lastly our present 'quasi-uniform' convergence corresponds to the possibility of securing the fundamental inequality by the inequalities $|x-\xi| <$ some $\delta(\epsilon)$ and $n >$ some $N(\epsilon)$. It represents an order of choice

$$\epsilon, \quad \delta, \quad N.$$

That the continuity in x, say, has this sort of 'quasi-uniformity' in

† Actually (29) leaves $f(x,y)$ undefined at any $(0,a_n)$, but it is convergent there as $x \to 0$, and we may therefore extend the definition to secure continuity.

‡ This is possible since $a_n \to 0$.

y near (ξ, η), if $f(x, y)$ is continuous in (x, y) there, is soon seen. For, by the general condition of convergence (18),

$$|f(x_1, y_1) - f(x_2, y_2)| < \epsilon$$

throughout some interval

$$|x - \xi| < \delta_1(\xi, \eta, \epsilon), \qquad |y - \eta| < \delta_2(\xi, \eta, \epsilon),$$

and therefore, in particular,

$$|f(\xi, y) - f(x, y)| < \epsilon$$

throughout this interval, which establishes the 'quasi-uniformity' in y of the continuity in x.

It is unprofitable to pursue these ideas with greater persistence. Enough has been said to show the distressing variety of possibilities in the convergence and continuity of functions of many variables.

11. Continuity throughout a closed region

We can generalize for functions of many variables continuous throughout a *closed region* the theorems of this chapter on functions of a single variable continuous throughout a closed interval.

For a function $f(x, y)$ we shall have the following theorems:

(30) *If $f(x, y)$ is continuous throughout a closed region,*

(i) *it is uniformly continuous over the region;*

(ii) *it is bounded in the region;*

(iii) *it attains in the region a greatest and a least value;*

(iv) *it attains in the region every value lying between $f(\xi_1, \eta_1)$ and $f(\xi_2, \eta_2)$, where (ξ_1, η_1) and (ξ_2, η_2) are points of the region;*

(v) *more particularly, it attains every such value on every continuous curve which joins (ξ_1, η_1), (ξ_2, η_2) and lies within the region.*

The necessary modifications in the argument will be sufficiently shown by a proof of (ii): that $f(x, y)$ *is bounded in a closed region throughout which it is continuous.*

For simplicity take the closed region to be the closed interval

$$a \leqslant x \leqslant a', \qquad b \leqslant y \leqslant b', \tag{31}$$

and suppose, if possible, that $f(x, y)$ is unbounded in this interval. Now, in particular, $f(x, b)$ is a continuous function of x (only) in the interval (a, a'), and therefore, by theorem (1) of this chapter, is bounded in the interval. We can therefore consider the numbers η such that $f(x, y)$ is bounded in the closed interval

$$a \leqslant x \leqslant a', \qquad b \leqslant y \leqslant \eta.$$

For, by what we have just said, $b \leqslant \eta$. Again, on the supposition that

$f(x, y)$ is unbounded in (31), we have $\eta < b'$. The set $\{\eta\}$ is thus bounded above, and we can consider η_0 its upper bound (or its greatest number). Evidently

$$b \leqslant \eta_0 \leqslant b'.$$

Then, by definition of η_0, the function $f(x, y)$ is bounded in any interval

$$a \leqslant x \leqslant a', \qquad b \leqslant y \leqslant \eta_0 - \epsilon$$

and is unbounded in any interval

$$a \leqslant x \leqslant a', \qquad b \leqslant y \leqslant \eta_0 + \epsilon',$$

where ϵ, ϵ' are any numbers, positive or zero, that, of course, satisfy the conditions

$$b \leqslant \eta_0 - \epsilon, \qquad b' \geqslant \eta_0 + \epsilon'$$

but are otherwise arbitrary, and in particular may be made as small as we please. It follows that $f(x, y)$ is unbounded in any interval

$$a \leqslant x \leqslant a', \qquad \eta_0 - \epsilon \leqslant y \leqslant \eta_0 + \epsilon'. \tag{32}$$

Now (a, η_0) is a point of the interval (31), so that $f(x, y)$ is continuous there and in particular

$$|f(x, y) - f(a, \eta_0)| < \tfrac{1}{2}|f(a, \eta_0)|,$$

if (x, y) lies in some interval

$$a \leqslant x \leqslant a+\delta, \qquad \eta_0 - \epsilon_1 \leqslant y \leqslant \eta_0 + \epsilon_1', \tag{33}$$

i.e. $f(x, y)$ is bounded in this interval. Fix ϵ_1, ϵ_1' and consider the set of numbers $\{\delta\}$ such that $f(x, y)$ is bounded in the interval (33). Since $f(x, y)$ is unbounded in such an interval as (32), we have $\delta < a'$, and the set has therefore an upper bound or greatest number δ_0 lying in the interval (a, a') and such that $f(x, y)$ is bounded in any interval

$$a \leqslant x \leqslant a+\delta_0-\epsilon_2, \qquad \eta_0 - \epsilon_1 \leqslant y \leqslant \eta_0 + \epsilon_1'$$

and unbounded in any interval

$$a \leqslant x \leqslant a+\delta_0+\epsilon_2', \qquad \eta_0 - \epsilon_1 \leqslant y \leqslant \eta_0 + \epsilon_1',$$

where ϵ_2, ϵ_2' are any numbers, positive or zero such that

$$0 \leqslant \delta_0 - \epsilon_2, \qquad a+\delta_0+\epsilon_2' \leqslant a'$$

but otherwise are arbitrary.

Hence $f(x, y)$ is unbounded in any interval

$$a+\delta_0-\epsilon_2 \leqslant x \leqslant a+\delta_0+\epsilon_2', \qquad \eta_0 - \epsilon_1 \leqslant y \leqslant \eta_0 + \epsilon_1'. \tag{34}$$

But $(a+\delta_0, \eta_0)$ is a point of the interval (31), so that $f(x, y)$ is continuous there and hence, in particular,

$$|f(x, y) - f(a+\delta_0, \eta_0)| < \tfrac{1}{2}|f(a+\delta_0, \eta_0)|,$$

if (x, y) lies in some interval

$$a+\delta_0-\epsilon_3 \leqslant x \leqslant a+\delta_0+\epsilon_3, \qquad \eta_0 - \epsilon_4 \leqslant y \leqslant \eta_0 + \epsilon_4;$$

in other words, $f(x,y)$ is bounded in this interval. But this is such an interval as (34), in which, as we have deduced from our original supposition, $f(x,y)$ is unbounded. Hence this supposition that $f(x,y)$ is unbounded in the interval (31) is shown to be untenable, and we have therefore proved that $f(x,y)$ is bounded in the interval (31), i.e. we have established the theorem (30) (ii).

The proof of (30) (v) follows immediately from (12), for, if the continuous curve be

$$x = x(t), \qquad y = y(t),$$

then, by (16), $f\{x(t), y(t)\}$, the value of the function on the curve, is a continuous function of the single variable t.

WORKED EXAMPLE

Show that a function $f(x)$ which, in a given interval (a,b), possesses either of the properties

(i) *it attains its greatest and least values for any closed sub interval (a',b') of (a,b) at least once in the sub-interval,*

(ii) *it attains at least once in any sub-interval (a',b') of (a,b) every value between $f(a'), f(b')$;*

does not necessarily possess the other.

Show further that a function which possesses both these properties in (a,b) is not necessarily continuous throughout (a,b).

To prove that (i) does not imply (ii) it is sufficient to consider the function defined as follows:

$$\left.\begin{aligned} f(x) &= 0 \qquad (x \text{ rational}) \\ &= 1 \qquad (x \text{ irrational}) \end{aligned}\right\}$$

For, since any interval (a',b') always contains both rational and irrational points, the greatest and least values of $f(x)$ in the interval are respectively 1 and 0, which are actually attained in the interval. But, by definition, the function does not attain any value between 0, 1, and therefore (ii) does not hold, if a, b are one rational, the other irrational.

To prove that (ii) does not imply (i) we may consider such a function as

$$\left.\begin{aligned} f(x) &= (1-x^2)\sin(x^{-2}) \qquad (x \neq 0) \\ f(0) &= 0 \end{aligned}\right\}$$

in the interval $(-1,1)$. The function is continuous except in the neighbourhood of $x = 0$. It therefore attains in (a',b') every value between $f(a'), f(b')$, if a', b' have the same sign. If a', b' have opposite signs, suppose that $|a'| \geq |b'|$, so that the interval $(a', -b')$ is part of the interval (a',b'). Then $f(x)$ assumes in $(a', -b')$ every value between $f(a'), f(-b')$. But $f(-b') = f(b')$ and thus $f(x)$ attains in (a',b') every value between $f(a'), f(b')$.

Finally, for an interval $(a',0)$, we may suppose $a' > 0$ and take n so great that

$$(n\pi)^{-\frac{1}{2}} < a'.$$

Thus the interval $(a',0)$ contains the interval $\{a', (n\pi)^{-\frac{1}{2}}\}$. But in this latter interval $f(x)$ is continuous and so attains every value between $f(a')$, $f\{(n\pi)^{-\frac{1}{2}}\}$, i.e. between $f(a'), f(0)$. Thus $f(x)$ has the property (ii).

Again 1 is a superior bound of $f(x)$ in $(-1, 1)$ for

$$1-x^2 < 1, \qquad \sin(x^{-2}) \leqslant 1 \qquad (x \neq 0),$$

i.e. $$f(x) < 1 \ (x \neq 0) \quad \text{and} \quad f(0) = 0.$$

But given any positive ϵ we can choose x in $(-1, 1)$ such that

$$f(x) > 1-\epsilon.$$

For choose a positive integer n so great that

$$(2n\pi+\tfrac{1}{2}\pi)^{-1} < \epsilon$$

and take $$x^2 = (2n\pi+\tfrac{1}{2}\pi)^{-1}.$$

Then $$(1-x^2)\sin(x^{-2}) = 1-x^2 > 1-\epsilon.$$

Thus 1 is actually the upper bound of $f(x)$ in $(-1, 1)$ and is never attained, because, as we have shown, $f(x) < 1$. Thus $f(x)$ does not possess the property (i).

Lastly, to show that (i), (ii) in combination are not sufficient to secure continuity throughout (a, b), we may consider such a function as

$$\left.\begin{array}{ll} f(x) = \sin(x^{-1}) & (x \neq 0) \\ f(0) = 0 & \end{array}\right\}.$$

In any interval not containing the origin $f(x)$ is continuous and therefore possesses both the properties (i), (ii), In every interval containing the origin $f(x)$ runs through its full range of values, namely all values between ± 1 inclusive, which are its bounding values In any such interval (a', b'), then, the function attains its bounding values and also every value between $f(a')$, $f(b')$ It accordingly possesses both the properties (i), (ii), but, as we know, it is not continuous throughout any interval containing the origin

EXAMPLES III

1 Determine the upper and lower bounds, or the greatest and least values, where they exist, of the real functions

$$\{(x-a)(b-x)\}^{\frac{1}{2}}, \qquad (\tan x - x)^{\frac{1}{2}}, \qquad x-I(x),$$
$$\frac{x^2-y^2}{x^2+y^2}, \qquad \frac{x^2-y^2}{x^4+y^4}, \qquad \frac{xI(y)-yI(x)}{x-y}.$$

2 Prove that the following functions have single-valued inverses

$$y(x) = x-\tan x, \qquad x-\sin x, \qquad x+\cos x, \qquad x-\sin^2 x$$

3. Prove that a function which is continuous and has a single-valued inverse in a closed interval is monotonic in the interval.

4. Determine (with their domains of definition) the branches of the inverses of the following functions

$$y(x) = \sin^2 x, \qquad x-I(x),$$
$$y(x) = ax^2+2bx+c, \quad \text{separating the cases } ac \gtrless b^2,$$
$$\frac{\sin x}{x}+\frac{\sin 2x}{2x}+\frac{\sin 3x}{3x}+\ldots \quad \text{to } \infty,$$
$$\cos x+\frac{1}{2}\frac{\cos^3 x}{3}+\frac{1.3}{2.4}\frac{\cos^5 x}{5}+\ldots \quad \text{to } \infty,$$
$$\sin x+\frac{1}{2}\frac{\sin^3 x}{3}+\frac{1.3}{2.4}\frac{\sin^5 x}{5}+\ldots \quad \text{to } \infty,$$
$$(1-\tfrac{1}{2}\pi)\sin x+\frac{1}{2}\frac{\sin^3 x}{3}+\frac{1.3}{2.4}\frac{\sin^5 x}{5}+\ldots \quad \text{to } \infty.$$

Determine also the branch-points.

5. Show that, if $f(x)$ is continuous throughout a closed interval, so also is $\{|f(x)|+1\}^{-1}$. Hence deduce Theorem (1) from Theorem (2) in the form: 'a continuous function, *if bounded*, actually attains its extreme values.'

6. If, throughout a closed interval, $f(x)$ is continuous and never zero, deduce from Theorems (2) and (14) that $f(x)$ is of constant sign in the interval.

7. If $u_n(x) \to u(x)$ uniformly in (a, b) and if $|u_n(x)| < A$ also in (a, b), prove that

$$F\{u_n(x)\} \to F\{u(x)\} \quad \text{uniformly in } (a, b),$$

if $F(t)$ is continuous in the closed interval $(-A, A)$.

8. Of the two properties

(i) $f(x)$ attains at least once in every sub-interval (a', b') of (a, b) every value between $f(a'), f(b')$;

(ii) for every ξ in (a, b) a sequence $\{x_n\}$ can be chosen such that $x_n \to \xi$, $f(x_n) \to f(\xi)$;

show that (i) implies (ii), but that (ii) does not imply (i).

9. If $f(x)$ is continuous throughout (a, b) and if $a < x_1 < x_2 < \ldots < x_n < b$, prove that

$$(x_1-a)[f(x_1)-f(a)]+(x_2-x_1)[f(x_2)-f(x_1)]+\ldots+(b-x_n)[f(b)-f(x_n)]$$

is numerically less than ϵ, provided that x_1-a, $x_2-x_1, \ldots, b-x_n$ are all less than some $\delta(\epsilon)$.

10. Prove that a rational integral algebraic function $f(x, y)$ is everywhere continuous, and that a rational algebraic function $f(x, y)/g(x, y)$ is continuous except at the zeros of the denominator.

11. Discuss the existence, and, if it exist, determine the value of the double limit $(x, y \to 0)$ of the following functions:

$$\frac{ax^2+2hxy+by^2}{a'x^2+2h'xy+b'y^2}, \quad \frac{ax^3+3bx^2y+3cxy^2+dy^3}{px^2+2qxy+ry^2}, \quad \frac{x^2y^2}{x^2y^2+(x-y)^2}, \quad \frac{x^py^q}{x^2+y^2}\ (p, q > 0),$$

$\dfrac{f(x, y)}{g(x, y)}$, where f, g are homogeneous functions of degrees m, n,

$$\frac{x+y^2}{x^2+y}, \quad \frac{x^2+y^4}{x^4+y^2}, \quad \frac{(x^2+y^4)(x+y^2)}{x^4+y^2}, \quad \frac{(x^2+y^4)(x^2+y)}{x^4+y^2},$$

$$(x+y)\sin(x^{-1})\sin(y^{-1}), \quad \cot\{(x^2+y^2)^{-1}\}, \quad \coth\{(x^2+y^2)^{-1}\}, \quad \cot^{-1}[(x^2-y^2)^{-1}],$$

$$\coth^{-1}[(x^2+y^2)^{-1}], \quad \tan^{-1}\operatorname{cosec}(x+y), \quad \log\cos xy, \quad |x|^y, \quad |x|^{|y|},$$

$$\exp[y^{-1}\sin(x^{-1})], \quad \exp[-y^{-2}\operatorname{cosec}^2(x^{-1})].$$

12. Prove that $f(x, y)$ is continuous at $(0, 0)$, if $f(r\cos\theta, r\sin\theta)$ is continuous in r at $r = 0$, the continuity being uniform in θ. Prove also the converse.

13. If $s(n, x, y)$ is continuous in (x, y) and converges to $s(x, y)$ uniformly in (x, y) over some region, prove that $s(x, y)$ is continuous in (x, y) throughout the region.

14. If $f(x)$ is continuous throughout the interval (a, b) and if $M(\xi)$ denotes its greatest value in the sub-interval (a, ξ), prove that $M(x)$ is continuous in (a, b).

15. If $f(x)$ is continuous throughout (a, b) and if $M(\xi, \eta)$ denotes its greatest value in the sub-interval (ξ, η), show that $M(x, y)$ is continuous in (x, y) throughout the square $a \leqslant x, y \leqslant b$.

16. If $f(x, y)$ is continuous throughout $(a, b; a', b')$ and if $M(\eta)$ denote the greatest value of $f(x, \eta)$ in (a, a'), prove that $M(y)$ is continuous in (b, b').

17. If $f(x, y)$ is continuous throughout $(a, b; a', b')$ and if $M(\xi, \eta)$ denote its greatest value in $(a, b; \xi, \eta)$, prove that $M(x, y)$ is continuous in (x, y) throughout $(a, b; a', b')$.

18. If $M(\xi, \eta)$ denote the greatest of $f_1(\xi, \eta), \ldots, f_n(\xi, \eta)$, where $f_1(x, y), \ldots, f_n(x, y)$ are all continuous at $x = a$, $y = b$, show that $M(x, y)$ is also continuous at (a, b).

IV

THE DERIVED FUNCTION

1. Differentiability

In the last two chapters we have been chiefly concerned with continuous functions of a single variable, that is to say with functions $f(x)$ such that, for some or all values of ξ,

$$f(x)-f(\xi) \to 0, \quad \text{as} \quad x-\xi \to 0.$$

Now, although in this way $f(x)-f(\xi)$ and $x-\xi$ vanish simultaneously, it does not also follow that they vanish in a finite or even in a definite ratio. Thus write

$$f(x) \equiv x^{\frac{1}{3}}, \qquad \xi = 0. \tag{1}$$

Then
$$f(x)-f(\xi) = x^{\frac{1}{3}} \to 0, \quad \text{as} \quad x-\xi \to 0,$$

but
$$\frac{f(x)-f(\xi)}{x-\xi} = x^{-\frac{2}{3}},$$

which diverges as $x-\xi \to 0$.

Again, write
$$\left.\begin{aligned} f(x) &\equiv x\sin(x^{-1}) \qquad (x \neq 0) \\ f(0) &= 0 \end{aligned}\right\}, \tag{2}$$

and take $\xi = 0$.

Then
$$f(x)-f(\xi) = x\sin(x^{-1}) \to 0, \quad \text{as} \quad x-\xi \to 0,$$

since
$$|f(x)-f(\xi)| \leqslant |x-\xi|.$$

But
$$\frac{f(x)-f(\xi)}{x-\xi} = \sin(x^{-1}),$$

which oscillates near $x = \xi$.

A function $f(x)$, such that the ratio

$$\frac{f(x)-f(\xi)}{x-\xi} \tag{3}$$

converges as $x \to \xi\pm$, is said to be *differentiable* (in x) at $x = \xi$. The examples (1), (2) above show that

(4) *A function continuous at a point is not necessarily differentiable there.*

On the other hand we can prove that

(5) *A function differentiable at a point is necessarily continuous there.*

For, if $f(x)$ is differentiable at $x = \xi$, then, as $x \to \xi$, the ratio (3) converges, and $x-\xi$ converges to zero. Thus their product $f(x)-f(\xi)$ also converges to zero, i.e. $f(x)$ is continuous at $x = \xi$.

It is often convenient to replace the above symbols ξ, x by x, $x+\delta x$.

In this notation δx is the typical *increment* of x; similarly $\delta f(x)$ denotes $f(x+\delta x)-f(x)$, the corresponding increment of $f(x)$. The *incremental ratio* (3) may then be written
$$\frac{\delta f(x)}{\delta x}. \tag{6}$$
This ratio is a function of both x and δx. If we denote it by $A(x,\delta x)$, we have the identity
$$\delta f(x) \equiv A(x,\delta x)\,\delta x. \tag{7}$$
The condition that $f(x)$ be differentiable at $x = \xi$ is then the condition that $A(\xi,\delta x)$ converge as $\delta x \to 0$. It is sometimes of advantage to use the condition of differentiability in this form.

If $f(x)$ is differentiable at $x = \xi$, the limit to which $A(\xi,\delta x)$ converges is called the 'derivative' (in x) or the 'x-derivative' of $f(x)$ at $x = \xi$.

If $f(x)$ is differentiable at all points of some domain, the corresponding set of values of the derivative defines, for that domain, a new function of x which we call the 'derived' function.

The operation of forming the incremental ratio (6) and proceeding to the limit is known as 'differentiation' (in x).

The present chapter—and indeed the remainder of the book—is occupied with the study of differentiable functions and of their corresponding derived functions. In the earlier sections I discuss the differentiability and the derivatives of the elementary functions. I then develop certain fundamental properties of differentiable functions and their derived functions.

2. Notation

The accepted notations for the derivative or the derived function are many. We may write $f'(x)$ or $f_1(x)$, or simply f' or f_1 when the variable of differentiation is not in doubt. These are akin to the original fluxional notation $\dot{f}$ of Newton, still sometimes used when time is the independent variable, but to be discouraged on typographical grounds.

To Leibniz we owe the symbol
$$\frac{df}{dx}$$
designed to recall the form of the incremental ratio (6). It indicates, of course, the limit of the ratio of δf, δx, *not* the ratio of their limits, which is the meaningless fraction $0/0$. The symbol is, in conception, one and indivisible.

None the less, the consequences of the fractional form are algebraically inconvenient, and mathematicians have sought to split the symbol, by devising some analogy to the relation (7) in the form
$$df(x) = f'(x)\,dx. \tag{8}$$

Attempts to attach meanings to these symbols $df(x)$, dx have not always been free from equivocation. It was once the custom to describe them as 'infinitesimals', being, by definition, infinitely small and yet different from zero. These 'infinitesimals' were known as the 'differentials' of $f(x)$ and x respectively. Thus, in the relation (8), the derivative $f'(x)$ appears as the coefficient of the 'differential' dx, and for this reason has long borne the name of *differential coefficient*. It seems better to preserve the name but to bury the idea.

We can give a sounder meaning to 'differential', if we define

$$dx \equiv \delta x, \qquad df(x) \equiv f'(x)\,dx,$$

so that the 'differential' of x is identically the increment of x, while the 'differential' of $f(x)$ is a function, both of x and of its increment.

But the drawback now is that the two differentials have been given dissimilar definitions, so that, if we are thinking geometrically and x, y are coordinates in a plane, their differentials will be differently defined according as we approach them *via* the function $x(y)$ or *via* the function $y(x)$. This, of course, cuts at the root of the symmetry in x, y which the 'differential' notation is specially designed to secure.

I shall presently make my own attempt to legitimize the equation (8). Meanwhile, we may observe that there is one unobjectionable splitting of Leibniz's symbol. We may write it

$$\frac{d}{dx}f(x),$$

where now the symbol d/dx denotes the *operation* of differentiating applied to $f(x)$. This operation is also frequently denoted by the symbol D_x, or simply by D if the variable of differentiation is obvious or unimportant.

3. Change of variable of differentiation

As we have seen in chapter II (16), a continuous function of a continuous function is itself continuous. It is an equally important theorem that a differentiable function of a differentiable function is itself differentiable. More precisely

(9) *If $y(x)$ is differentiable in x at $x = \xi$ and $f(y)$ is differentiable in y at $y = y(\xi)$, then $f\{y(x)\}$ is differentiable in x at $x = \xi$, and the derivative is*

$$f'(y)y'(x),$$

where $f'(y)$ is the y-derivative of $f(y)$ and $y'(x)$ is the x-derivative of $y(x)$.

Since $y(x)$ is differentiable at $x = \xi$, it is also continuous there, and

hence $|\delta y| < \epsilon$ is secured by $|\delta x| <$ some $\eta(\epsilon, \xi)$. Since $f(y)$ is differentiable at $y = y(\xi)$, we have, by (7),

$$\delta f(y) = A(y, \delta y)\, \delta y,$$

where $$A(y, \delta y) \to f'(y), \quad \text{as} \quad \delta y \to 0.$$

Thus, if $|\delta x| < \eta$, $$\frac{\delta f(y)}{\delta x} = A(y, \delta y)\, \frac{\delta y}{\delta x}.$$

Since $y(x)$ is differentiable, we may take the limit $\delta x \to 0$ on the right, and therefore also the consequential limit $\delta y \to 0$. This gives the desired formula

$$\frac{d}{dx} f(y) = f'(y) y'(x). \tag{10}$$

This formula enables us to pass from a variable y to a variable x of differentiation. We may write it as an identity of differential operators in the form

$$\frac{d}{dx} = y'(x)\, \frac{d}{dy}, \tag{11}$$

or even in the form

$$\frac{d}{dy} = \frac{d}{y'(x)\, dx}. \tag{11'}$$

If we write (10) in the form

$$D_x f(y) = f'(y)\, D_x y,$$

it is seen to be true for any variable x of differentiation, provided only that $y(x)$ is differentiable. We may thus ignore x and write simply

$$Df(y) = f'(y)\, Dy. \tag{12}$$

This equation is of the same symbolic form as the 'equation of differentials' (8), i.e.

$$df(x) = f'(x)\, dx.$$

We can thus give a reasonable meaning to this latter equation by defining d as the operator of differentiation with regard to an *arbitrary or unspecified parameter* (granted of course that x is differentiable in this parameter) or, what comes to the same thing, by writing Leibniz's symbol for the derivative in the form

$$\frac{Dy}{Dx},$$

where again D denotes differentiation with regard to any suitable unspecified parameter.

Implying this interpretation I shall make free use of the equation of differentials, writing d or D indifferently. With this interpretation equation (11′), for instance, appears as an immediate consequence of

$$dy = y'(x)\, dx.$$

With (9) we may associate the following theorem on the differentiation of inverse functions:

(13) *Wherever $y(x)$ has a non-zero derivative $y'(x)$, the inverse function $x = x(y)$ is differentiable, if it is continuous, and the derivative is given by*

$$x'(y) = 1/y'(x).$$

From the differentiability of $y(x)$ we have

$$\delta y = A(x, \delta x)\, \delta x,$$

where $A(x, \delta x) \to y'(x)$ as $\delta x \to 0$. We may write this as

$$\delta x = \{A(x, \delta x)\}^{-1}\, \delta y.$$

Since the inverse function is continuous, $\delta y \to 0$ secures $\delta x \to 0$, which again secures the convergence of $\{A(x, \delta x)\}^{-1}$, if $y'(x) \neq 0$. Dividing by δy and taking this limit $\delta y \to 0$, we at once have

$$x'(y) = 1/y'(x).$$

As we have seen in chapter III §§ 5, 6, a continuous inverse can generally be secured to a continuous function through the device of branch-points. The same device therefore, in general, secures a differentiable inverse to a differentiable function, zeros of the derivative being excluded.

4. Differentiation of sums and products

Let $u(x)$, $v(x)$ be two functions differentiable in x at $x = \xi$. It follows at once from the meaning of the symbol δ that

$$\delta(u \pm v) = \delta u \pm \delta v$$

and so
$$\frac{\delta(u \pm v)}{\delta x} = \frac{\delta u}{\delta x} \pm \frac{\delta v}{\delta x}.$$

The limit of a sum (difference) is the sum (difference) of the limits, and so the limit on the right exists as $\delta x \to 0$, and is $D_x u \pm D_x v$. Hence the limit on the left exists and has this value. It is enough to write

(14) $$D(u \pm v) = Du \pm Dv.$$

For this implies that $D(u \pm v)$ exists whenever Du, Dv separately exist, i.e. that $u(x) \pm v(x)$ is differentiable at $x = \xi$, if $u(x)$, $v(x)$ are each differentiable there. Moreover, we need not specify x, since the result is true for any parameter in which u, v are differentiable.

Again, for a product we have

$$\delta(uv) = (u+\delta u)(v+\delta v) - uv = u\, \delta v + (v+\delta v)\, \delta u,$$

and so
$$\frac{\delta(uv)}{\delta x} = u\frac{\delta v}{\delta x} + (v+\delta v)\frac{\delta u}{\delta x}.$$

On the right, as $\delta x \to 0$, both $\delta v/\delta x$ and $\delta u/\delta x$ converge, and $v+\delta v \to v$, since v, being differentiable, is also continuous. Hence the limit on the right exists and is $u\, D_x v + v\, D_x u$. We may therefore write

$$D(uv) = u\, Dv + v\, Du. \tag{15}$$

Save at a zero of $u(x)$ or $v(x)$ we may rewrite this as

$$\frac{D(uv)}{uv} = \frac{Du}{u} + \frac{Dv}{v}. \tag{16}$$

If $w(x)$ be a third function differentiable in x at $x = \xi$ and not vanishing there, we have, applying (6) to the product of uv and w,

$$\frac{D(uv.w)}{uv.w} = \frac{D(uv)}{uv} + \frac{Dw}{w}$$
$$= \frac{Du}{u} + \frac{Dv}{v} + \frac{Dw}{w},$$

by substitition from (16).

Hence, similarly, given n functions $u_1(x),\ldots, u_n(x)$ differentiable in x at $x = \xi$ and not zero there, we have

$$\frac{D(\Pi\, u_r)}{\Pi\, u_r} = \sum \frac{Du_r}{u_r}. \tag{17}$$

The differentiation of a product in this form is often called 'logarithmic differentiation' for a reason apparent later.

If we identify the n functions, (17) becomes

$$\frac{D(u^n)}{u^n} = n\frac{Du}{u}$$

and gives

$$D(u^n) = nu^{n-1}\, Du \quad (n \textit{ a positive integer}). \tag{18}$$

Since any polynomial can be resolved into sums and products of its arguments, it follows from (14), (15) that

(19) *Any polynomial whose arguments are differentiable in x at $x = \xi$ is itself differentiable in x at this point.*

Moreover, the form of its derivative is to be ascertained by sufficiently repeated applications of (14), (15).

5. Differentiation of a determinant

Important among such polynomials is the determinant whose elements are differentiable functions of x, say

$$\begin{vmatrix} u_1(x) & u_2(x) & u_3(x) & . \\ v_1(x) & v_2(x) & v_3(x) & . \\ w_1(x) & w_2(x) & w_3(x) & . \\ . & . & . & . \end{vmatrix}. \tag{20}$$

By definition it is an algebraic sum of products which we write symbolically

$$\Sigma \pm u_1(x)v_2(x)w_3(x).... \tag{21}$$

Now it follows from (15) or (17) that to differentiate a continued product we differentiate one factor at a time and add the results. In the derivative

$$D\{\Sigma \pm u_1(x)v_2(x)w_3(x)...\}$$

let us first collect those products

$$\{Du_r(x)\}v_s(x)w_t(x)...$$

in which the differentiated factor comes from the *first* row of the determinant. Agreeably with the notation (21) we may denote the algebraic sum of these products by

$$\Sigma \pm \{Du_1(x)\}v_2(x)w_3(x)....$$

Similarly, the algebraic sum of those products in which the differentiated factor comes from the second row is

$$\Sigma \pm u_1(x)\{Dv_2(x)\}w_3(x) ..$$

and so on. Hence

$$\begin{aligned} &D\{\Sigma \pm u_1(x)v_2(x)w_3(x) ..\} \\ &\quad = \Sigma \pm \{Du_1(x)\}v_2(x)w_3(x) .. + \Sigma \pm u_1(x)\{Dv_2(x)\}w_3(x)... + \\ &\qquad + \Sigma \pm u_1(x)v_2(x)\{Dw_3(x)\}... + ... \end{aligned} \tag{22}$$

Restoring the determinantal notation we have for the derivative of (20) the sum of determinants

$$\begin{vmatrix} Du_1(x) & Du_2(x) & Du_3(x) & . \\ v_1(x) & v_2(x) & v_3(x) & . \\ w_1(x) & w_2(x) & w_3(x) & . \\ . & . & . & . \end{vmatrix} + \begin{vmatrix} u_1(x) & u_2(x) & u_3(x) & . \\ Dv_1(x) & Dv_2(x) & Dv_3(x) & . \\ w_1(x) & w_2(x) & Dw_3(x) & . \\ . & . & . & . \end{vmatrix} +$$

$$+ \begin{vmatrix} u_1(x) & u_2(x) & u_3(x) & . \\ v_1(x) & v_2(x) & v_3(x) & . \\ Dw_1(x) & Dw_2(x) & Dw_3(x) & . \\ . & . & . & . \end{vmatrix} + ...$$

We may express this result more simply as a working rule:

(23) *To differentiate a determinant, differentiate each row (column) in turn and add the results.*

6. Differentiation of algebraic functions

Consider first the rational, integral algebraic functions. Since $\delta x/\delta x = 1$ and $\delta c = 0$, if c is a constant, we have the two trivial results

$$D_x x = 1, \quad Dc = 0, \tag{24}$$

and hence by (15)

(25) $$D(cy) = c\,Dy.$$

It follows from (19), (24) that

(26) *Any rational, integral algebraic function of x is everywhere differentiable in x.*

To differentiate any rational fractional function it is enough to be able to differentiate x^{-1}. For this, we have

$$\delta(x^{-1}) = \frac{1}{x+\delta x} - \frac{1}{x} = \frac{-\delta x}{x(x+\delta x)}.$$

Hence $$\frac{\delta(x^{-1})}{\delta x} = \frac{-1}{x(x+\delta x)},$$

and so, on taking the limit $\delta x \to 0$,

$$D_x(x^{-1}) = -x^{-2}. \tag{27}$$

The result is unintelligible, if $x = 0$, but holds for all other values of x. Moreover, we must at the start choose $|\delta x|$ less than $|x|$, so that $x+\delta x$ shall not pass through zero in the passage to the limit.

More generally, at a point where $v(x)$ is differentiable and not zero, we have by (9)

$$D\frac{1}{v} = -\frac{Dv}{v^2}, \tag{28}$$

and so from (15) $$D\frac{u}{v} = \frac{v\,Du - u\,Dv}{v^2}. \tag{29}$$

It follows from (26) that

(30) *Any rational, fractional algebraic function is differentiable, save at the zeros of its denominator.*

As we know, at these zeros the function is discontinuous and therefore *a priori* is not differentiable, by (5).

We may write (28) as $$\frac{Dv^{-1}}{v^{-1}} = -\frac{Dv}{v}$$

and so from (18) $$\frac{Dv^{-n}}{v^{-n}} = n\frac{Dv^{-1}}{v^{-1}} = -n\frac{Dv}{v},$$

or, in more usual form,

(31) $$Dv^{-n} = -nv^{-n-1}\,Dv.$$

Thus (18) extends also to a negative integral index.

For irrational algebraic functions it is similarly sufficient to differentiate x^p where p is not integral. In this case we make the convention, as in chapter II § 5, that x is positive or possibly zero.

Suppose firstly that x is positive, and take $|\delta x| < x$, so that $x+\delta x$ is also positive. Then by the exponential inequality

$$p(x+\delta x)^{p-1}\,\delta x \gtreqless (x+\delta x)^p - x^p \gtreqless px^{p-1}\,\delta x,$$

i.e.
$$\frac{(x+\delta x)^p - x^p}{\delta x}$$

lies between $p(x+\delta x)^{p-1}$ and px^{p-1}. But, by chapter II (20), x^{p-1} is continuous at a positive x and so, as $\delta x \to 0$,

$$p(x+\delta x)^{p-1} \to px^{p-1},$$

and therefore also
$$\frac{(x+\delta x)^p - x^p}{\delta x} \to px^{p-1}.$$

Thus x^p is differentiable, if x is positive, and its derivative is given by

(32) $$D_x x^p = px^{p-1}.$$

If x is zero, we have identically

$$\frac{\delta(x^p)}{\delta x} = (\delta x)^{p-1}.$$

This converges to 0, 1 if $p > 1$, $= 1$, but diverges if $p < 1$. Hence x^p is differentiable at $x = 0$, only if $p \geqslant 1$, and then (32) still holds.

Thus, by (31), (32), the formula (18) has been extended to any real index, and we may say, in general terms, that the formula holds for any value of x for which it has a meaning.

We observe that, if $0 < p < 1$, the function x^p is continuous at $x = 0$, by chapter II (20), but, as we have just seen, not differentiable there. This is in accordance with (4) above; in fact, we have already used the case $p = \frac{1}{3}$ in support of (4).

7. Differentiation of a^x, $\log x$

For the differentiation of a^x we need the fundamental limit†

$$\frac{a^x - 1}{x} \to \log_e a \quad \text{as} \quad x \to 0.$$

Then
$$\frac{a^{x+\delta x} - a^x}{\delta x} = \frac{a^x(a^{\delta x} - 1)}{\delta x} \to a^x \log_e a \quad \text{as} \quad \delta x \to 0.$$

Hence

(33) $$D_x a^x = a^x \log_e a$$

and, in particular,

(34) $$D_x e^x = e^x.$$

† Chrystal, *Algebra* 2 (1900), 79 (Cor. 4).

For the differentiation of $\log_a x$ we employ the fundamental limit

$$(1+x^{-1})^x \to e \quad \text{as} \quad x \to \infty.$$

Then

$$\frac{1}{\delta x}\{\log_a(x+\delta x)-\log_a x\}$$

$$= \log_a\left(1+\frac{\delta x}{x}\right)^{1/\delta x} = \frac{1}{x}\log_a\left\{\left(1+\frac{\delta x}{x}\right)^{x/\delta x}\right\}.$$

By chapter II (22), $\log x$ and therefore also $\log_a x$ is a continuous function, so that

$$\lim(\log) = \log(\lim).$$

Hence

$$\lim_{\delta x\to 0}\frac{1}{\delta x}\{\log_a(x+\delta x)-\log_a x\} = \frac{1}{x}\log_a e,$$

and so

(35) $$D_x \log_a x = \frac{1}{x}\log_a e,$$

and, in particular, to base e

(36) $$D_x \log x = \frac{1}{x}.$$

If we apply (9) it follows that, if Πu denote a product of factors $u_1(x)$, $u_2(x)$..., then

$$\frac{D\,\Pi u}{\Pi u} = D\log\,\Pi u$$

$$= D\sum\log u = \sum\frac{Du}{u}.$$

This gives an alternative proof of (17) and justifies the title of 'logarithmic differentiation'.

The hyperbolic functions are simple algebraic functions of e^x and so are at once differentiable by aid of (34). It is enough to state the derivatives

(37) $$D_x \sinh x = \cosh x, \qquad D_x \cosh x = \sinh x,$$

$$D_x \tanh x = \operatorname{sech}^2 x, \qquad D_x \coth x = -\operatorname{cosech}^2 x,$$

$$D_x \operatorname{sech} x = -\frac{\sinh x}{\cosh^2 x}, \qquad D_x \operatorname{cosech} x = -\frac{\cosh x}{\sinh^2 x}.$$

In like fashion the inverse hyperbolic functions are merely logarithms of algebraic functions. We have for their derivatives:

(38) $$D_x \sinh^{-1}x = D_x\log\{x+\sqrt{(x^2+1)}\} = (x^2+1)^{-\frac{1}{2}},$$

$$D_x \cosh^{-1}x = D_x\log\{x+\sqrt{(x^2-1)}\} = (x^2-1)^{-\frac{1}{2}},$$

$$D_x \tanh^{-1}x = D_x\,\tfrac{1}{2}\log\frac{1+x}{1-x} = (1-x^2)^{-1},$$

$$D_x \coth^{-1}x = D_x \tfrac{1}{2}\log\frac{x+1}{x-1} = -(x^2-1)^{-1},$$

$$D_x \operatorname{sech}^{-1}x = D_x \log\frac{1+\sqrt{(1-x^2)}}{x} = -x^{-1}(1-x^2)^{-\frac{1}{2}},$$

$$D_x \operatorname{cosech}^{-1}x = D_x \log\frac{1+\sqrt{(1+x^2)}}{x} = -x^{-1}(1+x^2)^{-\frac{1}{2}}.$$

8. Differentiation of the circular functions

The differentiation of the circular functions turns on the limit†

$$\frac{\sin x}{x} \to 1 \quad \text{as} \quad x \to 0.$$

For
$$\frac{\delta(\sin x)}{\delta x} = \frac{\sin(x+\delta x)-\sin x}{\delta x} = \frac{2\sin\frac{1}{2}\delta x\cos(x+\frac{1}{2}\delta x)}{\delta x}.$$

In this product, as $\delta x \to 0$, the factor $(\sin\frac{1}{2}\delta x)/\frac{1}{2}\delta x \to 1$, while the other factor converges to $\cos x$, since, by chapter II (23), the cosine is a continuous function. Hence

$$D_x \sin x = \cos x.$$

Thus also
$$\begin{aligned} D_x \cos x &= D_x \sin(\tfrac{1}{2}\pi-x) \\ &= \cos(\tfrac{1}{2}\pi-x)D_x(\tfrac{1}{2}\pi-x), \quad \text{by (9)}, \\ &= -\sin x. \end{aligned}$$

Again, by (29),

$$D_x \tan x = D_x\frac{\sin x}{\cos x} = \frac{\cos^2 x+\sin^2 x}{\cos^2 x} = \sec^2 x,$$

$$D_x \sec x = D_x\frac{1}{\cos x} = \frac{\sin x}{\cos^2 x}.$$

So
$$D_x \cot x = D_x \tan(\tfrac{1}{2}\pi-x) = -\operatorname{cosec}^2 x,$$

$$D_x \operatorname{cosec} x = D_x \sec(\tfrac{1}{2}\pi-x) = -\frac{\cos x}{\sin^2 x}.$$

Collecting these results we have for the derivatives of the circular functions:

(39)
$$\begin{aligned} D_x \sin x &= \cos x, & D_x \cos x &= -\sin x, \\ D_x \tan x &= \sec^2 x, & D_x \cot x &= -\operatorname{cosec}^2 x, \\ D_x \sec x &= \sin x/\cos^2 x & D_x \operatorname{cosec} x &= -\cos x/\sin^2 x \\ &= \tan x \sec x, & &= -\cot x \operatorname{cosec} x. \end{aligned}$$

It will be observed that a negative sign appears in the derivatives of the three 'co-functions' *cos*, *cot*, *cosec*, whereas the derivatives of the three 'direct' functions *sin*, *tan*, *sec* are positive in form. The inter-

† Cf. Hobson, *Plane Trigonometry* (1897), 123.

pretation is to be found in (42′) below: the direct functions increase, the co-functions decrease in the first quadrant.

We may use (13) to get the derivatives of the inverse circular functions. To the functions $\sin y$, $\tan y$, $\sec y$ with derivatives $\cos y$, $\sec^2 y$, $\tan y \sec y$ correspond the inverse functions

$$y = \sin^{-1}x, \qquad y = \tan^{-1}x, \qquad y = \sec^{-1}x$$

with derivatives

$$1/\cos y, \qquad 1/\sec^2 y, \qquad 1/\tan y \sec y,$$

i.e.
$$\frac{\pm 1}{\sqrt{(1-x^2)}}, \qquad \frac{1}{1+x^2}, \qquad \frac{\pm 1}{x\sqrt{(x^2-1)}}.$$

Since $\cos^{-1}x = \frac{1}{2}\pi - \sin^{-1}x$ and similarly for the inverses of the other co-functions, it is enough to record

$$(40) \qquad D_x \sin^{-1}x = \frac{\pm 1}{\sqrt{(1-x^2)}}, \qquad D_x \tan^{-1}x = \frac{1}{1+x^2},$$
$$D_x \sec^{-1}x = \frac{\pm 1}{x\sqrt{(x^2-1)}}.$$

To resolve the ambiguity of sign in the derivative of $\sin^{-1}x$ we remark that the sign is the sign of $\cos y$. Hence, if $\sin^{-1}x$ lies between $(2n \pm \frac{1}{2})\pi$, its derivative is $+(1-x^2)^{-\frac{1}{2}}$; if $\sin^{-1}x$ lies between $(2n+1 \pm \frac{1}{2})\pi$, its derivative is $-(1-x^2)^{-\frac{1}{2}}$. At the values $x = \pm 1$, where $\sin^{-1}x$ is $(n+\frac{1}{2})\pi$, the derivative does not exist.

We resolve the other ambiguity of sign in similar fashion. If $\sec^{-1}x$ lies between $2n\pi$, $(2n+1)\pi$, its derivative is $+x^{-1}(x^2-1)^{-\frac{1}{2}}$; if $\sec^{-1}x$ lies between $(2n-1)\pi$, $2n\pi$, its derivative is $-x^{-1}(x^2-1)^{-\frac{1}{2}}$. We must again exclude the values $x = \pm 1$.

For principal values of the inverse functions the positive signs are to be taken.

9. Functions differentiable throughout an interval. Maxima and minima

As we have seen at some length in chapter II, functions continuous throughout an interval have certain striking and sovereign properties. Functions differentiable throughout an interval form a sub-class possessing additional important properties in which the derived functions are in general also concerned. These functions and properties we now consider.

It must be emphasized at the outset that it is of the essence of differentiability that $\delta f/\delta x$ converge, as δx converges to zero through values *of either sign.*

Thus $f(x) \equiv |x|$ is not differentiable at $x = 0$, since there

$$\frac{\delta f}{\delta x} = \frac{|\delta x|}{\delta x} = \pm 1, \text{ according as } \delta x \gtrless 0.$$

Hence
$$\frac{\delta f}{\delta x} \to +1, -1, \quad \text{as} \quad \delta x \to 0+, 0-.$$

The existence of these left-handed and right-handed 'derivates', as they are called, has its own interest, but is insufficient to confer the title 'differentiable'.

An exception must, however, be made, for a function defined over a closed domain, at an end-point of this domain, since values of the function on one side only of the end-point are available. Thus $f(x) = x^{\frac{3}{2}}$ is differentiable at $x = 0$, since there

$$\frac{\delta f}{\delta x} = (\delta x)^{\frac{1}{2}} \to 0 \quad \text{as} \quad \delta x \to 0+,$$

while $(\delta x)^{\frac{1}{2}}$ has no meaning (from our present point of view) for negative δx.

We are led at once to a fundamental property of the derivative:

(41) *If a function is differentiable at a maximum or a minimum, the derivative vanishes there*

Suppose x a minimum of $f(x)$, so that throughout some interval $(x-\delta, x+\delta)$ the increment $\delta f(x)$ is positive and the incremental ratio $\delta f/\delta x$ has therefore the sign of δx.

We are given that $\delta f/\delta x$ converges as $\delta \to 0\pm$. Since $\delta f/\delta x$ is positive, as $\delta x \to 0+$, the limit can only be positive or zero; since $\delta f/\delta x$ is negative, as $\delta x \to 0-$, the limit can only be negative or zero. We draw the only possible conclusion that the limit $f'(x)$ is zero.

The maxima of $f(x)$ are the minima of $-f(x)$, and a like result therefore extends to maxima.

The function $|x|$ is, as we have seen, not differentiable at the minimum $x = 0$, and the right and left derivates there are respectively ± 1, and not zero.

The converse of (41), that $f(x)$ has a maximum or a minimum at every zero of $f'(x)$, is false. This is seen from the example $f(x) \equiv x^3$. Here $f'(0) = 0$, but the function is monotonic in any interval, since, if $x_1 \neq x_2$,

$$\frac{f(x_1)-f(x_2)}{x_1-x_2} = x_1^2+x_1x_2+x_2^2 > 0.$$

Hence there is neither maximum nor minimum at $x = 0$.

The zeros of the derivative are generally known as 'stationary' points of the function and the corresponding values of the function as 'stationary' values.

Consider now a function $f(x)$ differentiable throughout a closed interval (a, b) and therefore also continuous throughout the interval. For such a function we may consider in the interval the three categories of points not necessarily all different:

(i) points of greatest and least value,

(ii) maxima and minima,

(iii) stationary points.

By chapter III (2), there must be points (i); there may be points (ii), (iii). By (41) above, points (ii) must also be points (iii). Otherwise the categories may but need not overlap. We return to the more elaborate discussion of these topics in chapters IX, X.

10. Monotonic functions. Rolle's theorem

If $f'(x)$ exists throughout (a, b) without ever being zero, then $f(x)$ can have neither maximum nor minimum in the interval, and so, by chapter III (8), is monotonic there.

Thus the ratio
$$\frac{f(x_1)-f(x_2)}{x_1-x_2}.$$
has a constant sign in the interval. Suppose, for argument's sake, that it is always positive. Then, by taking the limit $x_2 \to x_1$, we have
$$f'(x_1) = \lim_{x_2 \to x_1} \frac{f(x_1)-f(x_2)}{x_1-x_2} \geqslant 0.$$

But $f'(x_1) = 0$ is excluded by hypothesis, and so $f'(x)$ shares the constant sign of $\{f(x_1)-f(x_2)\}/(x_1-x_2)$. Hence

(42) *If throughout an interval $f'(x)$ exists but does not vanish, then its sign is constant and $f(x)$ is monotonic in the interval.*

More precisely

(42′) *If $f'(x)$ is positive (negative) throughout an interval, then $f(x)$ is ascending (descending) monotonic in the interval.*

We often make use of (42′) in some such form as this:

if $\quad f'(x) > 0 \quad$ when $x > a$,

then also $\quad f(x)-f(a) > 0 \quad$ when $x > a$;

or again:

if $\quad f'(x) > g'(x) \quad$ when $x > a$, $\quad$ and if $\quad f(a) = g(a)$,

then $\quad f(x) > g(x) \quad$ when $x > a$.

This principle is of much service in proving inequalities. As an illustration let us establish the inequality already quoted in chapter I § 13

$$\operatorname{cosec}\tfrac{1}{2}x < \pi/x \quad (0 < x < \pi). \tag{43}$$

If $$f(x) \equiv x\operatorname{cosec}\tfrac{1}{2}x,$$

then $$f'(x) = \operatorname{cosec}\tfrac{1}{2}x - \tfrac{1}{2}x\cos\tfrac{1}{2}x\operatorname{cosec}^2\tfrac{1}{2}x$$
$$= \cos\tfrac{1}{2}x\operatorname{cosec}^2\tfrac{1}{2}x(\tan\tfrac{1}{2}x - \tfrac{1}{2}x).$$

Now this is positive in $)0,\pi)$, since $\tan\frac{1}{2}x > \frac{1}{2}x$ in $)0,\pi($ by the fundamental trigonometric inequality. Thus $f(x)$ is steadily increasing in $)0,\pi)$ and therefore

$$f(x) < f(\pi) \qquad (0 < x < \pi),$$

i.e. $$\operatorname{cosec}\tfrac{1}{2}x < \pi/x \qquad (0 < x < \pi),$$

which is (43).†

As a second illustration consider the three functions

$$\sinh x, \qquad x, \qquad \tanh x.$$

The functions are equal at $x = 0$. When $x \neq 0$, their derivatives have the order of magnitude $$\cosh x > 1 > \operatorname{sech}^2 x,$$

and therefore from our principle the original functions have, when x is positive, the same order of magnitude, namely,

$$\sinh x > x > \tanh x \qquad (x > 0).$$

We cannot, of course, apply the same principle to prove the corresponding inequalities for circular functions

$$\sin x < x < \tan x \qquad (0 < x < \tfrac{1}{2}\pi),$$

for that would be argument in a circle.

From (42) we at once deduce Rolle's theorem:

(44) *If $f'(x)$ exists throughout (a,b) and if $f(a) = f(b)$, then $f'(x)$ must vanish in the* open *interval $)a,b($;*

or in its more usual form:

(45) *In a region of differentiability $f'(x)$ has a zero between every two zeros of $f(x)$.*

For if $f'(x)$ does not vanish in (a,b), then $f(x)$ is monotonic, which contradicts $f(a) = f(b)$.

† If we write (43) in its more usual form

$$\sin x \geqslant \frac{2x}{\pi} \qquad (0 \leqslant x \leqslant \tfrac{1}{2}\pi),$$

we can deduce an inequality for the full interval $(0,\pi)$ in which $\sin x$ is positive, in the form

$$\sin x \geqslant \frac{2x(\pi - x)}{\pi^2}.$$

11. The theorem of the mean

If $f(x)$, $g(x)$ are two functions differentiable throughout (a, b) we can apply Rolle's theorem to the function

$$\begin{vmatrix} f(x) & g(x) & 1 \\ f(a) & g(a) & 1 \\ f(b) & g(b) & 1 \end{vmatrix},$$

which vanishes at $x = a$, b. Its derivative

$$\begin{vmatrix} f'(x) & g'(x) & 0 \\ f(a) & g(a) & 1 \\ f(b) & g(b) & 1 \end{vmatrix}$$

therefore vanishes at some point ξ of the interval $)a, b($. We have accordingly

$$[g(a)-g(b)]f'(\xi) = [f(a)-f(b)]g'(\xi), \tag{46}$$

which we write in the more convenient fractional form

$$\frac{f(a)-f(b)}{g(a)-g(b)} = \frac{f'(\xi)}{g'(\xi)}, \tag{47}$$

provided that neither denominator vanishes.

If $g(a)-g(b)$ vanish, it is clear from (46) that either $f(a)-f(b)$ vanishes also, in which case the equation is trivial, or else $g'(\xi)$ vanishes, which brings us back to Rolle's theorem. We thus lose nothing by supposing that neither $f(a)-f(b)$ nor $g(a)-g(b)$ vanishes.

If one only of $f'(\xi)$, $g'(\xi)$ vanishes, then either (47) or its reciprocal

$$\frac{g(a)-g(b)}{f(a)-f(b)} = \frac{g'(\xi)}{f'(\xi)}$$

is valid. If $f'(\xi)$, $g'(\xi)$ both vanish, the relation (46) again becomes trivial. We must therefore exclude the case in which $f'(x)$, $g'(x)$ have a common zero in $)a, b($, for in such a case we might well expect that ξ will be this zero.

In a rather more convenient notation we replace a, b by $a+h, a$. Then ξ lying between $a+h, a$ becomes $a+\theta h$, where $0 < \theta < 1$. We thus restate (47) as the 'theorem of the mean':

(48) *If, throughout* $(a, a+h)$, $f'(x)$, $g'(x)$ *exist and have no common zero, then*

$$\frac{f(a+h)-f(a)}{g(a+h)-g(a)} = \frac{f'(a+\theta h)}{g'(a+\theta h)}$$

(or else its reciprocal form) holds for some θ *in* $)0, 1($.

We should observe that θ can always be found in the *open* interval $)0, 1($. This is important in some applications.

The above gives the 'fractional' formula of the mean. If we take $g(x) \equiv x$, we have the 'linear' formula:

(49) *If $f(x)$ is differentiable throughout $(a, a+h)$, then*

$$f(a+h)-f(a) = hf'(a+\theta h)$$

for some θ in $)0, 1($.

It is to be noted that (48) is more precise than the form to be obtained by applying (49) successively to $f(x)$ and $g(x)$ and taking the ratio. For then we get on the right-hand side

$$\frac{f'(a+\theta h)}{g'(a+\theta_1 h)},$$

where θ, θ_1 agree as regards lying in the interval $)0, 1($ but are otherwise not necessarily the same. It must always be borne in mind that θ in (48) or (49) is essentially $\theta(a, h)$, defined in fact by the formula itself. The operative clause of (48) or (49) thus lies in the restriction $0 < \theta < 1$; these theorems of the mean, in other words, are disguised inequalities and not really statements of equality at all.

We can use (49) to give the following wider form of (42′) above:

(50) *If $f'(x)$ is never negative (positive) throughout an interval, then, as x increases, $f(x)$ never decreases (increases) in the interval: in other words $f(x)$ is monotonous (i.e. monotonic or stationary) in the interval.*

For, if x_1, x_2 are any two points of the interval, then

$$\frac{f(x_1)-f(x_2)}{x_1-x_2} = \text{some } f'(\xi),$$

where ξ lies between x_1, x_2, i.e. in the interval. Thus if $f'(x)$ is never negative in the interval, then $f(x_1)-f(x_2)$ has the sign of x_1-x_2 or is zero. Thus, as x increases, $f(x)$ increases or is stationary, i.e. never decreases. Thus (50) is proved.

In such an interval $f(x)$ can have neither maximum nor minimum, for if ξ were a maximum, we should have, for some x_1, x_2,

$$f(x_1) < f(\xi) > f(x_2),$$

where

$$x_1 < \xi < x_2,$$

so that

$$\frac{f(x_1)-f(\xi)}{x_1-\xi}, \qquad \frac{f(\xi)-f(x_2)}{\xi-x_2}$$

are one positive, one negative. Thus

(51) *If in a certain interval $f'(x)$ never changes sign, then $f(x)$ can have neither maxima nor minima in the interval.*

From (49) we deduce a theorem of great importance:

(52) *If $f'(x)$ vanishes throughout an interval, then $f(x)$ is constant throughout the interval.*

If $f'(x)$ vanishes throughout (a, b), it vanishes throughout any sub-interval $(a, a+h)$, and so every $f'(a+\theta h)$ is zero. Thus by (49)

$$f(a+h) = f(a),$$

i.e. $f(x) = f(a)$ throughout (a, b). It follows that

(52′) *Two functions which have the same derived function (throughout an interval) differ at most (throughout the interval) by an additive constant.*

This theorem is fundamental in the Integral Calculus and in the theory of Differential Equations: we can say that almost every differential equation at present completely soluble is reducible to the elementary differential equation

$$Df(x) = 0,$$

which, as we have just seen, has the solution

$$f(x) = \text{constant}.$$

It is instructive to compare the two formulae (7) and (49) for the increment of a function. In similar notation they are

$$(7) \quad \delta f(x) = A(x, \delta x)\,\delta x,$$

where $A(x, \delta x) \to f'(x)$ as $\delta x \to 0$;

$$(49) \quad \delta f(x) = f'(x+\theta\,\delta x)\,\delta x,$$

where $0 < \theta < 1$.

For (7), it is sufficient that $f'(x)$ exist only *at* the value; for (49), $f'(x)$ must exist throughout the interval $(x, x+\delta x)$. On the other hand, in (7) we may be inconvenienced by the presence of the unknown function $A(x, \delta x)$.

If $f'(x)$ exists in $(x, x+\delta x)$ and is also continuous at x, the two formulae are the same, for then $f'(x+\theta\,\delta x) \to f'(x)$ as $\delta x \to 0$. But, as we are now to see, the existence of the derived function near a point does not ensure its continuity at the point.

12. Discontinuity of the derived function

Consider the function

$$\left.\begin{aligned} f(x) &\equiv x^2\sin(x^{-1}) \qquad (x \neq 0) \\ f(0) &= 0 \end{aligned}\right\}. \tag{53}$$

Elsewhere than at the origin the function is differentiable by the rules of §§ 4–8 above, and we have in fact

$$f'(x) = 2x\sin(x^{-1}) - \cos(x^{-1}).$$

At the origin $$\frac{\delta f(x)}{\delta x} = \delta x \sin\{(\delta x)^{-1}\} \to 0 \quad \text{as} \quad \delta x \to 0.$$

Hence $$f'(0) = 0.$$

Now $x = 0$ is a point of continuity of $2x\sin(x^{-1})$, but a discontinuity (a point of finite oscillation) of $\cos(x^{-1})$ and therefore also of $f'(x)$. Hence

(54) *The derived function may exist in the neighbourhood of a point, and yet not be continuous at the point.*

On the other hand, it is worthy of note that a derived function, if it exists throughout an interval, shares some of the properties of a function continuous throughout the interval. In the first place

(55) *If $f'(x)$ exists near $x = \xi$, we can always find sequences of approach along which $f'(x)$ converges to $f'(\xi)$.*

For by the theorem of the mean (49) we have throughout the neighbourhood in question
$$\frac{f(x)-f(\xi)}{x-\xi} = f'(X),$$
where $X = X(x,\xi)$ and lies between x, ξ. Now choose any sequence $\{x_n\}$ converging to ξ. This gives the corresponding sequence $\{X_n\}$, which converges also to ξ, since X_n lies always between x_n and ξ. Since $f'(\xi)$ exists, the sequence of incremental ratios on the left converges to $f'(\xi)$, and therefore also the sequence $\{f'(X_n)\}$ on the right converges to $f'(\xi)$. Evidently we have as wide a choice for the sequence of approach $\{X_n\}$ as we have for the sequence $\{x_n\}$.

Again

(56) *If $f'(x)$ exists throughout (a,b), it assumes therein every value between $f'(a)$, $f'(b)$.*

For let k be a number lying between $f'(a)$, $f'(b)$ and consider the function
$$\phi(x) = f(x) - kx.$$
Its derivative $\phi'(x) = f'(x)-k$ exists throughout (a,b) and therefore, by (42), if it does not vanish in the interval, its sign is constant throughout the interval, and in particular at the end-points, i.e.
$$f'(a)-k, \qquad f'(b)-k$$
have the same sign. This is contrary to hypothesis and therefore $f'(x)-k$ must vanish somewhere in the interval, which proves the proposition.

Less precisely, we can restate (42) in the terms of chapter III (11):

(57) *A derived function can change sign only by passing through zero.*

We may ask if these two properties (55), (56) are independent. An example shows that (55) does not imply (56), for define the function

$$\left.\begin{aligned} f(x) &\equiv 0 \qquad (x \text{ irrational}) \\ f(x) &\equiv 1 \qquad (x \text{ rational}) \end{aligned}\right\}. \tag{58}$$

Since we can always approach an irrational number through a sequence of irrational numbers and can always approach a rational number through a sequence of rational numbers, it follows that we can always find a sequence of approach along which $f(x)$ converges to $f(\xi)$, although the function never assumes any value between 0, 1.

On the other hand, (56) *does* imply (55). For suppose that in every sub-interval (a, b) of some given interval, $f(x)$ assumes every value between $f(a)$ and $f(b)$, then, given any ξ in the interval and any positive ϵ, η, we can find $x(\epsilon, \eta)$ such that simultaneously

$$|x-\xi| < \epsilon, \qquad |f(x)-f(\xi)| < \eta. \tag{59}$$

To prove this take any ξ_1 lying between $\xi \pm \epsilon$. If $f(\xi_1)$ at the same time lies between $f(\xi) \pm \eta$, we have done what we promised. If not, then one of $f(\xi) \pm \frac{1}{2}\eta$ lies between $f(\xi)$ and $f(\xi_1)$, and we can therefore find x in (ξ, ξ_1) such that $f(x)$ is this one of $f(\xi) \pm \frac{1}{2}\eta$. Accordingly we have

$$|x-\xi| < |\xi_1-\xi| < \epsilon, \qquad |f(x)-f(\xi)| = \tfrac{1}{2}\eta < \eta,$$

which is (59).

Now choose two sequences $\{\epsilon_n\}$, $\{\eta_n\}$ both converging to zero and consider the corresponding sequence $x_n \equiv x(\epsilon_n, \eta_n)$. Then

$$|x_n-\xi| < \epsilon_n, \qquad |f(x_n)-f(\xi)| < \eta_n,$$

and the sequences $\{x_n\}$, $\{f(x_n)\}$ converge respectively to $\xi, f(\xi)$.

13. The bounds and infinities of the derived function

We may put into the scale *against* the derived function the following particulars in which it does not behave like a continuous function:

(60) *If a derived function exists throughout an interval, it is not necessarily bounded there;*

(61) *If a derived function exists and has an upper (lower) bound in an interval, it need not attain that upper (lower) bound in the interval.*

To prove (60) it is enough to exhibit the function

$$\left.\begin{aligned} f(x) &\equiv x^2\sin(x^{-2}) \qquad (x \neq 0) \\ f(0) &= 0 \end{aligned}\right\}. \tag{62}$$

Then $$f'(x) = 2x\sin(x^{-2}) - 2x^{-1}\cos(x^{-2}) \qquad (x \neq 0),$$
$$f'(0) = 0.$$

By considering, for instance, values $x = (n\pi)^{-\frac{1}{2}}$, it is seen that $f'(x)$ exists but is unbounded in the neighbourhood of the origin.

To prove (61) consider the function

$$\left.\begin{aligned} f(x) &= x^3\sin^2(x^{-2}) \qquad (x \neq 0) \\ f(0) &= 0 \end{aligned}\right\}. \tag{63}$$

Here $$f'(x) = 3x^2\sin^2(x^{-2}) - 2\sin(2x^{-2}) \qquad (x \neq 0),$$
$$f'(0) = 0.$$

In the neighbourhood of the origin $f'(x)$ has a lower bound -2, for, if $x \neq 0$, $\quad f'(x)+2 = 3x^2\sin^2(x^{-2}) + 2\{1-\sin(2x^{-2})\}.$

This is never negative and converges to zero, as x converges to zero through the set of values

$$x = \frac{2}{(2n+1)\sqrt{\pi}},$$

where n is a positive integer. But $f'(x)$ never attains this lower bound, for $f'(0) = 0$ and, if $x \neq 0$, $f'(x)+2 = 0$, only if simultaneously

$$\sin(x^{-2}) = 0, \qquad \sin(2x^{-2}) = 1,$$

which are incompatible conditions.

The infinities of a function and of its derived function are connected by the proposition:

(64) *An infinity of a function is also an infinity of the derived function, provided the latter exists throughout some half-neighbourhood of the infinity.*

If $x = \xi$ is such an infinity, $f(x)$ is discontinuous there and so $f'(\xi)$ cannot exist. We therefore suppose that $f'(x)$ exists throughout the half-neighbourhood $)\xi, a)$.

If x be a point of this interval, we have, by (49), that

$$f(x) - f(a) = (x-a)f'(X),$$

where X lies between x, a. By taking x sufficiently near ξ we can make $f(x)$, and therefore $f'(X)$, as great as we please. Moreover, a is still at our disposal. Hence in any interval $(\xi, \xi+h)$, however small, we can find x such that $f'(x)$ is as great as we please, i.e. ξ is an infinity of $f'(x)$.

The condition that the derivative exist throughout a half-neighbourhood is essential to the proposition; for consider the function

$$f(x) \text{ is the greatest integer in } x^{-1}. \tag{65}$$

The function is discontinuous at the countable set of points $x = \pm n^{-1}$,

where n is a positive integer. Elsewhere the derivative exists and is bounded, being everywhere zero. Nevertheless the origin is an infinity of $f(x)$.

14. Differentiation of a limit-function

In chapter II § 7 we considered the continuity of a function defined as the limit of a sequence of continuous functions. Here we similarly consider the differentiability of a function defined as the limit of a sequence of differentiable functions. More precisely we consider the differentiability at $x = \xi$ of the limit-function $s(x)$ defined by the sequence of functions $s(n, x)$ themselves differentiable at $x = \xi$. We go further and inquire whether, when this derivative $s'(\xi)$ does exist, it is actually given as the limit of the sequence of derivatives $s'(n, \xi)$.

Since the differentiability of $s(x)$ at $x = \xi$ is only another name for the convergence of $\{s(x)-s(\xi)\}/(x-\xi)$ as $x \to \xi$, it is clear that differentiability, like convergence and continuity, need not survive passage to the limit. Denoting by $s_1(x)$ the limit of $s'(n, x)$, when it exists, we can give examples of the three types of cases of failure:

(i) $s'(\xi)$ exists but not $s_1(\xi)$;
(ii) $s_1(\xi)$ exists but not $s'(\xi)$;
(iii) $s'(\xi)$, $s_1(\xi)$ both exist but are not equal.

For (i), it is enough to consider the logarithmic series

$$s(n, x) \equiv x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \ldots \text{to } n \text{ terms}, \tag{66}$$

so that $s(x)$ is $\log(1+x)$ and $s'(1)$ exists. But $s'(n, 1) = 1-1+1-\ldots$. This is an oscillating series and $s_1(1)$ does not exist.

For (ii), consider the sequence

$$s(n, x) \equiv (1+nx^2)^{-1}. \tag{67}$$

Then $s'(n, x) = -2nx(1+nx^2)^{-2}$, $s_1(0) = 0$. But $s(x) = 0$ $(x \neq 0)$, $s(0) = 1$, i.e. $s(x)$ is discontinuous at $x = 0$ and $s'(0)$ does not exist.

For (iii), consider the sequence

$$s(n, x) \equiv x(1+nx^2)^{-1}. \tag{68}$$

Then $s'(n, x) = (1-nx^2)(1+nx^2)^{-2}$, $s_1(0) = 1$. But $s(x) = 0$, $s'(x) = 0$, and so

$$s'(0) \neq s_1(0).$$

A condition *sufficient* both for the existence of $s'(\xi)$ and for the equality $s'(\xi) = s_1(\xi)$ is given by the theorem:

(69) *If in the neighbourhood of $x = \xi$ the differentiated sequence $s'(n, x)$ converges uniformly to $s_1(x)$, then $s(x)$, the limit-function of the sequence $s(n, x)$, is differentiable at $x = \xi$ and moreover $s'(\xi) = s_1(\xi)$.*

By the theorem of the mean (49)

$$\left|\frac{s(n+p,x)-s(n+p,\xi)}{x-\xi}-\frac{s(n,x)-s(n,\xi)}{x-\xi}\right| = |s'(n+p,X)-s'(n,X)|,$$

where X lies somewhere in $)x,\xi($. But, by the uniform convergence of $s'(n,x)$,

$$|s'(n+p,X)-s'(n,X)| < \tfrac{1}{3}\epsilon$$

for every X in some interval $(\xi-\delta,\xi+\delta)$, every $n >$ some $N(\epsilon,\xi)$ and every p. Hence

$$\left|\frac{s(n+p,x)-s(n+p,\xi)}{x-\xi}-\frac{s(n,x)-s(n,\xi)}{x-\xi}\right| < \tfrac{1}{3}\epsilon$$

for every x in $(\xi-\delta,\xi+\delta)$, every $n > N(\epsilon,\xi)$ and every p. Since p is unrestricted we may take the limit $p\to\infty$. This gives

$$\left|\frac{s(x)-s(\xi)}{x-\xi}-\frac{s(n,x)-s(n,\xi)}{x-\xi}\right| \leqslant \tfrac{1}{3}\epsilon \tag{70}$$

under the same conditions.

Again, since $s'(n,\xi)\to s_1(\xi)$,

$$|s'(n,\xi)-s_1(\xi)| < \tfrac{1}{3}\epsilon, \tag{71}$$

if $n >$ some $N'(\epsilon,\xi)$ Fix $n >$ both N, N'. Then, since $s'(n,\xi)$ is the derivative of $s(n,x)$ at $x=\xi$,

$$\left|\frac{s(n,x)-s(n,\xi)}{x-\xi}-s'(n,\xi)\right| < \tfrac{1}{3}\epsilon, \tag{72}$$

if $|x-\xi| <$ some $\delta(\epsilon,n,\xi)$.

By combining the three inequalities (70), (71), (72), we have

$$\left|\frac{s(x)-s(\xi)}{x-\xi}-s_1(\xi)\right| < \epsilon.$$

Since now n, as well as p, has disappeared, the inequality is secured provided only that $|x-\xi| <$ some $\delta(\epsilon,\xi)$. In other words $s'(\xi)=s_1(\xi)$, and the theorem is proved.

The uniform convergence of $s'(n,x)$, although sufficient, is not necessary. This is shown by such an example as

$$s(n,x) \equiv n^{-1}(1+n^2x^2)^{-1}. \tag{73}$$

Here $s'(n,x) = -2nx(1+n^2x^2)^{-2}$, $s_1(x)=0$. But $s'(n,x)$ does not converge uniformly to $s_1(x)$ near $x=0$, for, if $x, n\to 0,\infty$ respectively, so that $xn=1$, then $s'(n,n^{-1}) = -\tfrac{1}{2}$. Nevertheless $s(x)=0$, $s'(0)=0=s_1(0)$.

If we seek conditions that shall be both necessary and sufficient, we must inquire, in the spirit of chapter II § 8, for conditions necessary and

sufficient that $\dfrac{s(x)-s(\xi)}{x-\xi}$, the limit-function of $\dfrac{s(n,x)-s(n,\xi)}{x-\xi}$, converge as $x \to \xi$ and converge to the limit $s_1(\xi)$.

We may remark that the uniform convergence of $s'(n,x)$ near $x = \xi$ is insufficient to secure even the ordinary convergence of $s(n,x)$ near ξ. Thus, for example, $s(n,x) \equiv \log(n+x^2)$ diverges everywhere, but its derivative

$$s'(n,x) = \frac{2x}{x^2+n}$$

is comparable with x/n and converges uniformly to zero in any finite interval. On the other hand, by (70), the uniform convergence of $s'(n,x)$ near $x = \xi$ is sufficient to secure the uniform convergence of $s(n,x)-s(n,\xi)$ near $x = \xi$, and so, if some one $s(n,\xi)$ converges, then in addition every $s(n,x)$ converges uniformly near $x = \xi$. On the other hand, the uniform convergence of $s(n,x)$ does not imply that of $s'(n,x)$, for

$$s(n,x) \equiv x(1+nx^2)^{-1}$$

converges uniformly in any finite interval, since $1+nx^2 \geq 2|x|\sqrt{n}$ and so

$$|s(n,x)| < \tfrac{1}{2}n^{-\frac{1}{2}}.$$

But $s'(n,x) = (1-nx^2)(1+nx^2)^{-2}$ does not converge uniformly near $x = 0$. For

$$s_1(x) = 0 \quad (x \neq 0), \qquad s_1(0) = 1.$$

Thus $s'(n,x)-s_1(x) = \frac{12}{25}$, if $x = \frac{1}{2}n^{-\frac{1}{2}}$, and so $s'(n,x)$ does not converge uniformly near $x = 0$.

If we apply (69) to a power-series $\sum a_n x^n$, the differentiated power-series is $\sum na_n x^{n-1}$. The intervals of convergence of these two series are given† (in the usual notation) by

$$\rho = \underline{\lim}\,|a_n|^{-1/n}, \qquad \rho' = \underline{\lim}\,|na_n|^{-1/n}.$$

Since $\lim n^{-1/n} = 1$, these two intervals are the same. Hence the differentiated series converges uniformly over any closed interval within the interval of convergence of the original series. Thus, by (69),

(74) *A function represented by a power-series is differentiable within the interval of convergence and the derived function is represented by the differentiated series.*

15. Some pathological specimens

It is possible by means of infinite series to define functions which exhibit remarkable behaviour *qua* differentiation. Consider firstly the function

$$f(x) \equiv \sum \frac{1}{q^3}\left(x-\frac{p}{q}\right)^2 \sin\left\{\left(x-\frac{p}{q}\right)^{-1}\right\}, \tag{75}$$

† Cf. Bromwich, *Theory of Infinite Series* (1908), 28.

where the summation $\sum$ is taken over all positive proper fractions p/q that are in their lowest terms, and where, if x equals any p/q, the corresponding term is omitted from the summation. To fix the order of the summation let us first collect those terms which have the same q. There are at most $q-1$ positive proper fractions in lowest terms having q for denominator. This group of terms is thus together less than $x^2(q-1)/q^3$, and, if we now arrange in order of increasing q, the sum is comparable with $\sum q^{-2}$ and so is both absolutely convergent and uniformly convergent over any finite interval.

We have seen above† that the function

$$\left.\begin{aligned} \phi(x) &\equiv x^2\sin(x^{-1}) \qquad (x \neq 0) \\ \phi(0) &= 0 \end{aligned}\right\}$$

is everywhere differentiable, its derivative being given by

$$\begin{aligned} \phi'(x) &= 2x\sin(x^{-1})-\cos(x^{-1}) \qquad (x \neq 0), \\ \phi'(0) &= 0. \end{aligned}$$

The differentiated series is thus

$$f_1(x) = \sum \frac{1}{q^3}\left[2\left(x-\frac{p}{q}\right)\sin\left\{\left(x-\frac{p}{q}\right)^{-1}\right\}-\cos\left\{\left(x-\frac{p}{q}\right)^{-1}\right\}\right], \tag{76}$$

the term corresponding to p/q being again omitted, if $x = p/q$. This again is comparable for convergence with the convergent series $\sum q^{-2}$, and so is uniformly convergent over any finite interval. Hence, by (69), $f(x)$ is everywhere differentiable, and its derived function is $f_1(x)$.

Now $\phi'(x)$, as we saw in § 12 (53), is discontinuous at $x = 0$, but continuous everywhere else. Hence, except at rational points in $)0, 1($, the series (76) is a uniformly convergent series of continuous functions, i.e. it defines a continuous function. At a rational point $x = p/q$ in $)0, 1($, by a similar argument, $f_1(x)-\phi(x-p/q)$ is continuous, but $\phi(x-p/q)$ itself is discontinuous and therefore $f_1(x)$ is discontinuous.

Hence $f(x)$, as defined in (75), is *everywhere differentiable, but its derived function is discontinuous at every rational point in* $)0, 1($. It would not be difficult to define in similar fashion a function everywhere differentiable but with a derived function discontinuous at *all* rational points.

Remembering that $x\sin(x^{-1})$ is everywhere continuous and differentiable everywhere except at $x = 0$, we can construct a function

$$f(x) \equiv \sum \frac{1}{q^3}\left(x-\frac{p}{q}\right)\sin\left\{\left(x-\frac{p}{q}\right)^{-1}\right\} \tag{77}$$

which is everywhere continuous but is differentiable at no rational point

† § 12 (53).

in)0, 1(. But following Weierstrass, and others† we can go further and define functions *everywhere continuous but nowhere differentiable.* Such a function is

$$s(x) \equiv \sum_1^\infty a_n \sin(n!\,\pi x), \tag{78}$$

where $\sum a_n$ is a convergent series of positive constants. The series (78) is thus uniformly convergent over any finite interval and its terms are everywhere continuous. Hence also $s(x)$ is everywhere continuous.

To prove that, with suitable a_n, it is not differentiable at any given point ξ we proceed thus. Choose any positive integer n (for convenience divisible by 8). Write $$(n-1)!\,\xi = b+\epsilon,$$
where b is the integer *nearest* to $(n-1)!\,\xi$, so that ϵ may have either sign and $|\epsilon| \leqslant \frac{1}{2}$. Again write $$n\epsilon = c+\eta,$$
where, similarly, c is the integer nearest to $n\epsilon$ and $|\eta| \leqslant \frac{1}{2}$.

Then $|c+\eta| \leqslant \frac{1}{2}n$ and we can choose an integer c', namely one of $c\pm\frac{1}{4}n$, such that $$|c'-c| = \tfrac{1}{4}n, \text{ an even integer,}$$
and at the same time $$|c'+\eta| \leqslant \tfrac{1}{2}n.$$
Now, in the neighbourhood of ξ, take the point

$$x = \xi+\frac{c'-c}{n!},$$

so that

$$|x-\xi| = \frac{1}{4(n-1)!}.$$

Then the series for $f(x)$, $f(\xi)$ are the same after the $(n-1)$th term,

i.e.
$$f(x)-f(\xi) = \sum_{r=1}^{n-1} a_r\{\sin(r!\,\pi x) - \sin(r!\,\pi\xi)\}$$
$$= \sum_{r=1}^{n-1} 2a_r \sin\tfrac{1}{2}r!\,\pi(x-\xi)\cos\tfrac{1}{2}r!\,\pi(x+\xi).$$

Now we can make the first $n-2$ terms of this finite series less important than the last term u_{n-1}, for $\sum_1^{n-2} < \sum_1^{n-2} a_r\, r!\,\pi|x-\xi|$, since $r!\,\pi|x-\xi| < \frac{1}{2}\pi$,

i.e.
$$< \frac{\pi}{4(n-1)!}\sum_1^{n-2} a_r\, r!.$$

Take $a_r \equiv h^r/r!$, where $h > 1$. Then

$$\left|\sum_1^{n-2}\right| < \frac{\pi h^{n-1}}{4(h-1)(n-1)!}.$$

† See Hardy, *Trans. American Math. Soc.*, 17 (1916), 301–25, where further references are given; see also van der Waerden, *Math. Zeitschrift*, 32 (1930), 474–5 quoted in Titchmarsh, *The Theory of Functions* (1932), 353.

But $$u_{n-1} = a_{n-1}\left\{\sin\left(b+\frac{c'+\eta}{n}\right)\pi - \sin\left(b+\frac{c+\eta}{n}\right)\pi\right\}.$$

Since $|c'+\eta|$, $|c+\eta| < \frac{1}{2}n$, the angles $\left(b+\frac{c'+\eta}{n}\right)\pi$, $\left(b+\frac{c+\eta}{n}\right)\pi$ are in the same quadrant and differ by $\frac{1}{4}\pi$. Hence

$$|u_{n-1}| > a_{n-1}\left(1-\frac{1}{\sqrt{2}}\right),$$

i.e. $$> \frac{h^{n-1}}{(n-1)!}\left(1-\frac{1}{\sqrt{2}}\right).$$

Thus $$|f(x)-f(\xi)| > |u_{n-1}| - \left|\sum_1^{n-2} u_r\right|$$

i.e. $$> \frac{h^{n-1}}{(n-1)!}\left\{1-\frac{1}{\sqrt{2}}-\frac{\pi}{4(h-1)}\right\}.$$

Suppose h taken big enough to make the expression in $\{\quad\}$ positive.

Then $$\frac{f(x)-f(\xi)}{x-\xi} > 4h^{n-1}\left\{1-\frac{1}{\sqrt{2}}-\frac{\pi}{4(h-1)}\right\},$$

which diverges as $n \to \infty$. But $x \to \xi$ as $n \to \infty$. Hence the incremental ratio cannot converge as $x \to \xi$, i.e. $f(x)$ is not differentiable at any point ξ.

WORKED EXAMPLE

If $f(x)$ is continuous throughout (a,b) and if at every point ξ in (a,b)

$$\frac{f(\xi+h)-f(\xi-h)}{h} \to 0 \quad as \quad h \to 0,$$

then $f(x)$ is constant throughout (a,b).

It should first be remarked that the convergence, as $h \to 0$, of the special incremental ratio $$h^{-1}\{f(\xi+h)-f(\xi-h)\}$$

does not secure the existence of $f'(\xi)$. To see that, it is enough to consider the functions $$|x|, \qquad \sin^2(x^{-1})$$

at $x = 0$. Neither function is differentiable there, yet, since both functions are even functions, the special incremental ratio for each of them vanishes identically and so certainly converges as $h \to 0$.

We may see that the theorem is not necessarily true for a discontinuous function by considering the function

$$\left.\begin{aligned} f(x) &= 1 \qquad (x \neq 0) \\ f(0) &= 0 \end{aligned}\right\}. \tag{1}$$

Here, if $\xi \neq 0$, $\quad f(\xi+h)-f(\xi-h) = 0$, if $|h| < |\xi|$,

while $\quad f(h)-f(-h) = 0$ for every h.

Thus the limit of the special incremental ratio is everywhere zero, but the function is not constant in any interval which contains the origin.

To prove the theorem we state the condition of convergence in the form

$$|f(x+h)-f(x-h)| \leqslant h\epsilon, \quad \text{if} \quad 0 < h < \text{some } \eta(\epsilon, x), \tag{2}$$

where it is sufficient to consider h positive, since the expression on the left is an even function of h. I shall show that the numbers η are unrestricted, except, of course, by the condition that $x \pm h$ lie in the stated interval (a, b) For suppose, on the contrary, that for given x, ϵ the numbers η have a greatest number or an upper bound $\eta_0(\epsilon, x)$ Then

$$|f(x+h)-f(x-h)| \leqslant h\epsilon, \quad \text{if} \quad 0 < h < \eta_0(\epsilon, x), \tag{3}$$

$$|f(x+h')-f(x-h')| > h'\epsilon, \quad \text{for some } h' < \text{any } \eta_0+\epsilon', \tag{4}$$

where, of course $h' \geqslant \eta_0$, by (3).

Now, by (2),

$$|f(x+\eta_0+h)-f(x+\eta_0-h)| \leqslant h\epsilon, \quad \text{if} \quad 0 < h < \eta(\epsilon, x+\eta_0), \tag{5}$$

$$|f(x-\eta_0+h)-f(x-\eta_0-h)| \leqslant h\epsilon, \quad \text{if} \quad 0 < h < \eta(\epsilon, x-\eta_0) \tag{6}$$

Choose as ϵ' the smallest of η_0 and $\eta(\epsilon, x \pm \eta_0)$ Then, by (4), we can find a positive $h_1 < \epsilon' \leqslant \eta_0$ such that

$$|f(x+\eta_0+h_1)-f(x-\eta_0-h_1)| > (\eta_0+h_1)\epsilon \tag{7}$$

But by (5), (6)

$$|f(x+\eta_0+h_1)-f(x+\eta_0-h_1)| \leqslant h_1\epsilon, \quad \text{since} \quad 0 < h_1 < \eta(\epsilon, x+\eta_0),$$

$$|f(x-\eta_0+h_1)-f(x-\eta_0-h_1)| < h_1\epsilon, \quad \text{since} \quad 0 < h_1 < \eta(\epsilon, x-\eta_0),$$

and by (3)

$$|f(x_0+\eta_0-h_1)-f(x-\eta_0+h_1)| < (\eta_0-h_1)\epsilon, \quad \text{since} \quad 0 < \eta_0-h_1 < \eta_0$$

By addition of the last three inequalities we have

$$|f(x+\eta_0+h_1)-f(x-\eta_0-h_1)| \leqslant (\eta_0+h_1)\epsilon,$$

which contradicts (7)

The preceding argument is fallacious, however, if, in (4), h' is always η_0, i e if

$$|f(x+\eta_0)-f(x-\eta_0)| > \eta_0\epsilon > 0,$$

but

$$|f(x+h)-f(x-h)| \leqslant h\epsilon$$

in every open neighbourhood of $h = \eta_0$ We now require the condition that $f(x)$ be continuous, in virtue of which we can take the limit $h \to \eta_0$ This gives

$$|f(x+\eta_0)-f(x-\eta_0)| \leqslant \eta_0\epsilon,$$

and so rules out our assumption

The original hypothesis, then, that there is a greatest $\eta(\epsilon, x)$ is also disproved and we therefore have that

$$|f(x+h)-f(x-h)| \leqslant h\epsilon,$$

provided *only* that $x \pm h$ lie in (a, b) Thus ϵ is independent of x, h, and we may take the limit $\epsilon \to 0$, giving $\quad f(x+h) = f(x-h)$

for every h. In particular, if we write

$$x = h+c,$$

where c is a constant in (a, b), we get

$$f(x+2c) = f(c),$$

i.e $f(x)$ is constant throughout (a, b).

The essential argument of the proof shows that, even if continuity be not stipulated, the only discontinuous functions which satisfy the main condition of the theorem are those typified by the example (1) above, i.e. functions having discontinuities at a finite set of points $\xi_1,\ldots,\xi_n$ and such that at the discontinuity ξ_r

$$f(\xi_r+) = f(\xi_r-) \neq f(\xi_r).$$

It can also be seen that the theorem still holds, if the incremental ratio of the question be replaced by the more general ratio

$$h^{-1}\{f(x+ph)-f(x-qh)\},$$

where p, q are any two positive numbers *independent* of x, h.

EXAMPLES IV

1. Differentiate

$$\tan^{-1}\frac{a\sin x+b\cos x}{a\cos x-b\sin x},\qquad \frac{1}{1-n}\sin^{-1}(x^{1-n}),\qquad \sin^{-1}\frac{a+b\cos x}{b+a\cos x}\quad (a<b),$$

$$\log(\sec x+\tan x),\qquad \log\tan\tfrac{1}{2}x,\qquad \log\{\sqrt{(x+1)}+\sqrt{(x-1)}\},$$

$$\log\{\sqrt{(x+1)}+\sqrt{(x-1)}\}-\tfrac{1}{2}x\sqrt{(x^2-1)},$$

$$\frac{1}{\sqrt{(a-2)}}\tan^{-1}\frac{x\sqrt{(a-2)}}{x^2+1}-\frac{1}{\sqrt{(a+2)}}\tan^{-1}\frac{x\sqrt{(a+2)}}{x^2-1},$$

$$x\sqrt{(1-x^2)}\sin^{-1}x-\tfrac{1}{2}x^2+\tfrac{1}{2}(\sin^{-1}x)^2,$$

$$x\log(1+x^2)\tan^{-1}x+\log(1+x^2)+(\tan^{-1}x)^2-\tfrac{1}{2}\{\log(1+x^2)\}^2-2x\tan^{-1}x.$$

2. If a, b are positive, prove that

$$(1+a)\log(1+a)+(1+b)\log(1+b) < (1+a+b)\log(1+a+b),$$

and generally that, if every a_r is positive,

$$\sum_{r=1}^{n}(1+a_r)\log(1+a_r) < (1+\sum a_r)\log(1+\sum a_r).$$

3. If x is a positive acute angle, prove

(i) that x lies between $\sin x$ and $\sinh x$, and also between their arithmetic and geometric means;

(ii) that x lies between $\tan x$ and $\tanh x$ and is less than their arithmetic mean, their geometric mean, and their harmonic mean.

4. If x is positive and $\tan^{-1}x$ denotes the positive acute angle of given tangent, prove that $(\tan^{-1}x)/\tanh x$ is monotonic and that

$$\tan^{-1}x-\tfrac{1}{2}\pi\tanh x$$

is positive.

5. (i) In the fractional formula of the mean write

$$f(x) = 4x^3+6x^2-12x,\qquad g(x) = 3x^4+4x^3-6x^2,$$

so that $\dfrac{f'(x)}{g'(x)}=\dfrac{1}{x}$, $\dfrac{f(1)-f(0)}{g(1)-g(0)} = -2 \neq \dfrac{f'(x)}{g'(x)}$ in $(0, 1)$. Discuss this.

(ii) If $f'(x)$, $g'(x)$ exist throughout (a,b), and if $f(a)=f(b)$, $g(a)=g(b)$, show that $f'(x)/g'(x)$ assumes all real values in (a,b). [YOUNG.]

6. If the linear formula of the mean be applied, in turn, to the functions $\log x$ and e^x, determine the corresponding values of θ as functions of a, h. Obtain the inequalities

$$0 < [\log(1+x)]^{-1}-x^{-1} < 1,$$

$$0 < x^{-1}\log\left(\frac{e^x-1}{x}\right) < 1.$$

7. If $f(x)$ is differentiable throughout $(a-h, a+h)$, prove that

(i) $$f(a+h)-f(a-h) = h[f'(a+\theta_1 h)+f'(a-\theta_1 h)],$$

where $0 < \theta_1 < 1$;

(ii) $$f(a+h)-2f(a)+f(a-h) = h[f'(a+\theta_2 h)-f'(a-\theta_2 h)],$$

where $0 < \theta_2 < 1$.

8. If, in the fractional formula of the mean, we write for $f(x)$, $g(x)$, (i) x^2, x, (ii) $\sin x$, $\cos x$, (iii) e^x, e^{-x}, show that $a+\theta h$ is the arithmetic mean of $a, a+h$ and determine pairs of functions in which $a+\theta h$ is respectively the geometric and the harmonic mean of $a, a+h$.

9. In the linear formula of the mean determine the most general $f(x)$ such that $a+\theta h$ is (i) the arithmetic mean, (ii) the geometric mean, of $a, a+h$.

Show that, except for the trivial case of a linear function, we cannot choose $f(x)$ to make $a+\theta h$ identically the harmonic mean of $a, a+h$

10. In an interval throughout which $f'(x)$ exists it has opposite signs (or is zero) at consecutive zeros of $f(x)$.

11. Show that the derived function cannot have a break in value, i.e. show that $f'(\xi+), f'(\xi-), f'(\xi)$ must be equal, if they all exist.

12. If $f'(x)$ exists near $x = \xi$, show that its convergence as $x \to \xi$ is necessary and sufficient for the convergence as $h, k \to 0$ of the function of two variables

$$\frac{f(\xi+h)-f(\xi+k)}{h-k}.$$

13. Show that the existence of the limit

$$\lim_{h\to 0}\frac{f(\xi+h)-f(\xi-h)}{h}$$

is insufficient to establish the differentiability of $f(x)$ at $x = \xi$

More generally show that, if p, q are any two constants, the existence of the limit

$$\lim_{h\to 0}\frac{f(\xi+ph)-f(\xi+qh)}{h}$$

is insufficient to establish the differentiability of $f(x)$ at $x = \xi$

14. If in some domain of x

$$\left|\frac{f(x+h)-f(x)}{h}-f'(x)\right| < \epsilon$$

for $|h| < \eta(\epsilon)$ independent of x, we say that $f(x)$ is *differentiable uniformly* over the domain. Prove that the uniform differentiability of $f(x)$ over an interval is necessary and sufficient for the continuity of $f'(x)$ throughout the interval.

15. Discuss the differentiation at $x = 0$ of the sum functions of the series

(i) $$x\sin x+\tfrac{1}{2}x\sin 2x+\tfrac{1}{3}x\sin 3x+\ldots,$$

(ii) $$|x|\sin x+\tfrac{1}{2}|x|\sin 2x+\tfrac{1}{3}|x|\sin 3x+\ldots.$$

16. Prove that the 'denominator' function

$$\left.\begin{aligned} f(x) &= 0 \qquad (x \text{ irrational}) \\ f(x) &= 1/q \qquad (x \text{ rational} = p/q) \end{aligned}\right\}$$

is nowhere differentiable, although it is continuous at all irrational points.

17. Show that of the two functions

(i) $$\sum_{n=1}^{\infty} \frac{1}{n^2} \prod \sin\left\{\left(x-\frac{r}{n}\right)^{-1}\right\},$$

(ii) $$\sum_{n=1}^{\infty} \frac{\sin^2 \pi n x}{n^5} \prod \sin\left\{\left(x-\frac{r}{n}\right)^{-1}\right\},$$

where $\prod$ is taken for every r less than n and prime to it and at $x = r/n$ the factor $\sin\{(x-r/n)^{-1}\}$ is replaced by zero, (i) is discontinuous at all rational points in $)0, 1($; (ii) is everywhere differentiable, but its derivative is discontinuous at all rational points in $)0, 1($.

18. Prove that

$$f(x) = x^2 \sin\frac{1}{x} + \sum \frac{1}{p^3 q^3}\left\{(x-p/q)^2 \sin\frac{1}{x-p/q} + (x+p/q)^2 \sin\frac{1}{x+p/q}\right\},$$

where the term $(x-p/q)^2 \sin\{(x-p/q)^{-1}\}$ is omitted at $x = p/q$ and the corresponding term at $x = -p/q$, and the summation $\sum$ is taken over every interprime pair of positive integers p, q, is everywhere differentiable, but that its derived function is discontinuous at all rational points.

19. Show that the sum-functions of the infinite series

$$\sin x + \frac{\sin 2x}{2^2} + \frac{\sin 3x}{3^2} + \ldots, \quad \cos x + \frac{\cos 2x}{2^2} + \frac{\cos 3x}{3^2} + \ldots$$

are not differentiable at $x = 0$ and congruent points, but are differentiable at all other points.

More generally, if a_n is positive and steadily decreasing, prove that the sum-functions of

$$a_1 \sin x + \tfrac{1}{2}a_2 \sin 2x + \tfrac{1}{3}a_3 \sin 3x + \ldots \quad \text{to } \infty,$$

$$a_1 \cos x + \tfrac{1}{2}a_2 \cos 2x + \tfrac{1}{3}a_3 \cos 3x + \ldots \quad \text{to } \infty$$

are not differentiable at $x = 0$, if $\underline{\lim}\, n a_n > 0$.

20. If $s(n, x) \equiv a_1 x + \ldots + a_n x^n$ converges to $s(x)$ in the closed interval $(0, 1)$, and if we write

$$s(n, 1) = s_n, \qquad s(1) = s,$$

so that $s_n \to s$, show, in the notation of § 14, that $s'(1)$ exists, if

$$(s-s_1)+(s-s_2)+(s-s_3)+\ldots$$

converges, while $s_1(1)$ exists, if

$$a_1 + 2a_2 + 3a_3 + \ldots$$

converges, the differentiations at $x = 1$ being effected from below.

Prove that these two series both converge to the same limit or both diverge, if every a_n is positive.

If $a_n = (-)^n c_n$, where c_n converges steadily to zero, so that $\sum a_n$ is now a convergent alternating series, show that $\sum (s-s_n)$ is also a convergent alternating series, if $c_{n-1} - 2c_n + c_{n+1}$ is always positive, and that, in particular, this is the case if $c_n = f(n)$, where $f'(n)$ is ascending monotonic.

Consider the examples

$$c_n = \frac{1}{n}, \quad \frac{1}{\sqrt{n}}, \quad \frac{1}{n} + \frac{1}{n+1}, \quad \frac{1}{\sqrt{n}} + \frac{1}{\sqrt{(n+1)}},$$

showing that $s'(1)$ exists but not $s_1(1)$.

V

HIGHER DERIVATIVES

1. Definition and notation

In the previous chapter we have considered the existence and formation of the derivatives of one-valued functions of a single variable. In particular we found that the elementary functions are always differentiable save at exceptional points and that, moreover, their derived functions are also elementary functions. In the same way we found that a function represented by a power-series is differentiable throughout the interval of convergence with the possible exception of the end-points, the derived function being represented by the differentiated power-series. Most of the functions in common use are covered by these two types of function and we are therefore led to consider differentiation of the derived function itself.

If $f(x)$ is the original function, we now call $f'(x)$ its *first derivative* or *first derived function*. The derivative or derived function of $f'(x)$ is called the *second derivative* or *second derived function* of $f(x)$. We write it under any of the notations

$$f''(x), \quad f_2(x), \quad \frac{d^2f(x)}{dx^2}, \quad D^2f(x).$$

In like fashion by successive differentiation of derived functions it may be possible to proceed inductively up to an nth *derivative* or nth *derived function*, written† as

$$f^{(n)}(x), \quad f_n(x), \quad \frac{d^nf(x)}{dx^n}, \quad D^nf(x).$$

We call n the *order* of the derivative.

The existence of the nth derivative $f_n(\xi)$ means the existence of the limit

$$\lim_{h\to 0}\frac{f_{n-1}(\xi+h)-f_{n-1}(\xi)}{h},$$

and therefore presupposes the existence of the $(n-1)$th derivative at and near $x = \xi$. This, in turn, presupposes the existence at and near $x = \xi$ of the derivatives of all lower orders. If, then, we say that $f_n(\xi)$ exists, we must be understood as implying that all the derived functions $f_1(x), f_2(x),..., f_{n-1}(x)$ exist at and near $x = \xi$, and, in addition, that the last of them is differentiable at ξ. The others are, of course, differentiable

† The notation f_n is unorthodox but very convenient.

at and near ξ, and therefore, in particular, are continuous at and near ξ, while $f_{n-1}(x)$ is continuous at ξ.

More especially, if we say that the nth derived function $f_n(x)$ exists throughout an interval (a,b), we imply that every derived function $f_1(x), f_2(x), \ldots, f_{n-1}(x)$ of order less than n also exists throughout the interval, and is differentiable and therefore continuous throughout the interval.

2. Formulae for nth derivatives

The formulae given in chapter IV §§ 4–8 enable us, by successive differentiation, to obtain the nth derivative of any elementary function. A compact expression for the nth derivative as a function of n and x would be of great practical convenience, but in most cases is unobtainable. It is therefore worth while to collect such few formulae as are available.

The formulae of chapter IV § 4 for the differentiation of sums and products can be generalized without difficulty. The formula for the derivative of a sum or difference,

$$D(u \pm v) = Du \pm Dv,$$

leads at once to the corresponding formula for the nth derivative

$$D^n(u \pm v) = D^n u \pm D^n v. \tag{1}$$

The formula for a product,

$$D(uv) = u\,Dv + v\,Du,$$

has for its generalization *Leibniz's formula*

$$D^n(uv) = \sum_{r=0}^{n} (n\,!\,r) D^r u D^{n-r} v, \tag{2}$$

where $(n\,!\,r)$ *denotes*† *the binomial coefficient* $\dfrac{n!}{r!(n-r)!}$.

The proof comes most simply by induction. For assume that, for some n,

$$D^n(uv) = uD^n v + (n\,!\,1)DuD^{n-1}v + (n\,!\,2)D^2uD^{n-2}v + \ldots + (D^n u)v,$$

and differentiate. On the right we use the formula for the differentiation of a product and rearrange the terms. We then have

$$D^{n+1}(uv) = uD^{n+1}v + \{1+(n\,!\,1)\}DuD^n v + \{(n\,!\,1)+(n\,!\,2)\}D^2uD^{n-1}v + \\ + \ldots + (D^{n+1}u)v.$$

The coefficients on the right are those in the expansion of $(1+x)^{n+1}$, as can be recognized by expanding the identity

$$(1+x)^{n+1} = (1+x)(1+x)^n.$$

† Again an unorthodox but convenient notation.

We have therefore that

$$D^{n+1}(uv) = uD^{n+1}v+(n+1\,!\,1)DuD^nv+(n+1\,!\,2)D^2uD^{n-1}v+ +\ldots+(D^{n+1}u)v,$$

and so the formula holds for $n+1$. It holds also for $n = 1$, and is therefore established inductively.

It follows from (1), (2) that the sum and product of two functions $u(x)$, $v(x)$ can be differentiated n times at $x = \xi$, if the functions themselves can be differentiated n times at the point.

For the nth derivative of a continued product we apply (2) to successive partial products and so generalize it to give the formula

(3) $$D^n(u_1 u_2 \ldots u_m) = \sum \frac{n!}{a!\,b!\ldots k!} D^a u_1\, D^b u_2 \ldots D^k u_m,$$

where the coefficients are those in the multinomial expansion of

$$(x_1+x_2+\ldots+x_m)^n$$

and the summation $\sum$ is taken over all values of $a, b \ldots, k$ whose sum is n.

Passing to algebraic functions we have, by successive application of the formula

$$D(x-a)^p = p(x-a)^{p-1}$$

for diminishing values of p, that

(4) $$D^n(x-a)^p = p(p-1)\ldots(p-n+1)(x-a)^{p-n}.$$

This holds for all values of x, unless $p < n$, when we must exclude $x = a$.

If p is a positive integer not less than n, we can write (4) as

(5) $$D^n(x-a)^p = \frac{p!}{(p-n)!}(x-a)^{p-n}.$$

If p is a positive integer less than n, we have

$$D^n(x-a)^p = 0.$$

Hence, more generally,

(6) *A polynomial of degree n in x is annihilated by more than n differentiations.*

Conversely, we prove that

(7) *In an interval throughout which $f_n(x)$ vanishes $f(x)$ is a polynomial of degree less than n.*

For by chapter IV (52), since the derivative of $f_{n-1}(x)$ vanishes throughout the interval, we have therein

$$f_{n-1}(x) = c_0,$$

where c_0 is some constant. Hence the derivative of

$$f_{n-2}(x)-c_0 x$$

vanishes throughout the interval and so $f_{n-2}(x)-c_0 x$ is some constant c_1 in the interval. Proceeding in this way we have at length

$$f(x) = \frac{c_0 x^{n-1}}{(n-1)!}+\frac{c_1 x^{n-2}}{(n-2)!}+\ldots+c_{n-1}$$

throughout the interval, where the c's are constant and may be some or all zero.†

Passing to the elementary transcendental functions we have

$$D\log(x-a) = (x-a)^{-1},$$

and hence by (4)

(8) $$D^n\log(x-a) = (-)^{n-1}(n-1)!(x-a)^{-n}.$$

Again by repeated application of the formula

$$De^{ax} = ae^{ax},$$

we have

(9) $$D^n e^{ax} = a^n e^{ax}.$$

Also $$D(\sin x, \cos x) = (\cos x, -\sin x),$$

and so $$D^2(\sin x, \cos x) = (-\sin x, -\cos x).$$

Hence, more generally,

(10) $$D^{2n}(\sin x, \cos x) = (-)^n(\sin x, \cos x).$$

We may combine (9), (10) under a more general formula by considering the nth derivative of

$$f(x) \equiv e^{ax}\sin(bx+c).$$

Now

$$\begin{aligned} Df(x) &= e^{ax}\{a\sin(bx+c)+b\cos(bx+c)\} \\ &= ke^{ax}\sin(bx+c+\alpha), \end{aligned}$$

where k, α are new constants defined by the relations

$$a = k\cos\alpha, \qquad b = k\sin\alpha.$$

Thus a single differentiation of $f(x)$ multiplies by k and changes c into $c+\alpha$. The form of the function is otherwise unchanged and the rule therefore holds for successive differentiations. Hence at length

(11) $$D^n e^{ax}\sin(bx+c) = k^n e^{ax}\sin(bx+c+n\alpha),$$

where $$a = k\cos\alpha, \quad b = k\sin\alpha.$$

Sometimes the determination of $f_n(x)$ is facilitated by a preliminary transformation of $f(x)$ into the form of a sum.

Thus $$\cos^3 x = \tfrac{3}{4}\cos x+\tfrac{1}{4}\cos 3x$$

and so $$D^n\cos^3 x = \tfrac{1}{4}\{3\cos(x+\tfrac{1}{2}n\pi)+3^n\cos(3x+\tfrac{1}{2}n\pi)\}.$$

† Cf. also § 5 below.

Again $$\frac{1}{x^2+x^3} = \frac{1}{1+x}-\frac{1}{x}+\frac{1}{x^2},$$

and so $$D^n\frac{1}{x^2+x^3} = (-)^n n!\left\{\frac{1}{(1+x)^{n+1}}-\frac{1}{x^{n+1}}+\frac{n+1}{x^{n+2}}\right\}.$$

3. The Wronskian

A determinant of variable terms in which each row (column) is formed by differentiating the preceding row (column) is termed a *Wronskian*.† Such a determinant is

$$\begin{vmatrix} f_1(x) & f_2(x) & . & f_n(x) \\ Df_1(x) & Df_2(x) & . & Df_n(x) \\ . & . & . & . \\ D^{n-1}f_1(x) & D^{n-1}f_2(x) & . & D^{n-1}f_n(x) \end{vmatrix}, \tag{12}$$

where $f_1(x), f_2(x),\ldots,f_n(x)$ are now n independent functions, the suffixes no longer denoting differentiations with respect to x. We may conveniently write the above in the shortened symbolic form

$$(1, D,\ldots, D^{n-1})\,|f_1(x)\, f_2(x)\ldots f_n(x)| \tag{13}$$

or, even more compactly,

$$(1, D)^{n-1}|f_1(x)\, f_2(x)\ldots f_n(x)|. \tag{14}$$

By chapter IV (23) the derivative of a determinant is the sum of the determinants obtained by differentiating the given determinant one row at a time. If the given determinant is a Wronskian, all the differentiated determinants save one vanish by having two rows identical. The one determinant which survives is that obtained by differentiating the last row of the Wronskian. In other words, the derivative of (13) is the single determinant which agreeably with our notation we may write

$$(1, D,\ldots, D^{n-2}, D^n)\,|f_1(x)\, f_2(x)\ldots f_n(x)| \tag{15}$$

The characteristic property of a Wronskian is given by the theorem:

(16) *If throughout some interval the n functions $f_1(x), f_2(x),\ldots,f_n(x)$ have derivatives of order $n-1$ and obey the linear relation*

$$a_1 f_1(x)+a_2 f_2(x)+\ldots+a_n f_n(x) = 0,$$

the a's being constants not all zero, then their Wronskian

$$(1, D)^{n-1}|f_1(x)\, f_2(x)\ldots f_n(x)|$$

also vanishes throughout the interval.

† After Wronski who introduced them in the early nineteenth century.

The proof is immediate, for by $n-1$ differentiations of the linear relation we have the n equations

$$\begin{aligned} a_1 f_1(x)+a_2 f_2(x)+...+a_n f_n(x) &= 0 \\ a_1 Df_1(x)+a_2 Df_2(x)+...+a_n Df_n(x) &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_1 D^{n-1}f_1(x)+a_2 D^{n-1}f_2(x)+...+a_n D^{n-1}f_n(x) &= 0. \end{aligned}$$

Since the a's do not all vanish, the determinant of the equations must vanish, i.e.

$$(1,D)^{n-1}|f_1(x)f_2(x)...f_n(x)| = 0$$

throughout the interval.

A set of functions obeying such a linear relation may be called a *linearly connected* or *linearly dependent* set. A set of functions which obeys no such relation may similarly be called *linearly independent.* Thus the vanishing of the Wronskian throughout an interval is a *necessary* condition for linear *dependence* in that interval, and, equivalently, a *sufficient* condition for linear *independence* in an interval is that the Wronskian do not vanish throughout the interval. So much follows from (16). The converse of (16) is, however, not true: that is to say, the vanishing of the Wronskian throughout an interval is not sufficient to secure linear dependence in that interval. An example will show this. Write

$$f_1(x) = I(x)\sin^2 \pi x, \tag{17}$$

$$f_2(x) = \sin^2 \pi x,$$

where $I(x)$ denotes the greatest integer in x.

In an interval $)n, n+1($ between two consecutive integers, $I(x)$ is constant and

$$f_1'(x) = I(x)\pi \sin 2\pi x. \tag{18}$$

At an integral value $x = n$ we find, by actually evaluating the limit of the incremental ratio as $x \to n\pm$, that

$$f_1'(n) = 0.$$

Thus the formula (18) holds for every x, integers included. Since $f_2'(x) = \pi \sin 2\pi x$, for every x, we have throughout any interval that

$$f_1(x)f_2'(x)-f_2(x)f_1'(x) = 0,$$

i.e.

$$(1,D)|f_1(x)f_2(x)| = 0.$$

But the functions $f_1(x), f_2(x)$ do not satisfy a linear identity

$$a_1 f_1(x)+a_2 f_2(x) = 0,$$

for this would require $f_1(x)/f_2(x)$, i.e. $I(x)$ to be constant.

The two functions are, however, linearly connected in any interval $)n, n+1($ which does not include a zero of the functions. We might even

say very loosely that the functions are everywhere connected by a linear relation, but that the constants of the relation change, as we cross a zero $x = n$ of the functions, a proposition which becomes of little value, if the zeros are infinitely numerous in any neighbourhood.

If, however, the Wronskian of $f_1(x)$, $f_2(x)$ vanishes throughout an interval which contains no zeros of at least one of the functions, say $f_1(x)$, then we at once deduce that the functions are linearly connected throughout the interval. For, since $f_1(x) \neq 0$ in the interval, we may write the Wronskian condition in the form

$$\frac{f_2'(x)\,f_1(x)-f_2(x)\,f_1'(x)}{f_1^2(x)} = 0,$$

i.e. $$D\{f_2(x)/f_1(x)\} = 0,$$

i.e. $$f_2(x) = c\,f_1(x), \quad \text{where } c \text{ is constant.}$$

For n functions $f_1(x), f_2(x),\ldots, f_n(x)$, we similarly stipulate that there be at least one set of $n-1$ of them, say,

$$f_1(x),\ldots, f_{r-1}(x), f_{r+1}(x),\ldots, f_n(x)$$

whose Wronskian does not vanish in the interval.† The converse of (16) now becomes:

(19) *If the Wronskian*

$$(1, D)^{n-1}|f_1(x)\,f_2(x)\ldots f_n(x)|$$

vanishes throughout an interval, but one of the Wronskians

$$(1, D)^{n-2}||f_1(x)f_2(x)\ldots f_n(x)||$$

vanishes nowhere in the interval, then the functions $f_1(x), f_2(x),\ldots, f_n(x)$ are linearly connected throughout the interval.

Write $$\Delta \equiv (1, D)^{n-1}|f_1(x)f_2(x)\ldots f_n(x)|$$

and $$\Delta_1,\ldots, \Delta_n \equiv (1, D)^{n-2}||f_1(x)f_2(x)\ldots f_n(x)||.$$

These latter determinants are minors of Δ corresponding to terms in the last row. Since they are Wronskians, their derivatives $D\Delta_1,\ldots, D\Delta_n$ are

$$(1,\ldots, D^{n-3}, D^{n-1})||f_1(x)f_2(x)\ldots f_n(x)||.$$

These again are minors of Δ corresponding to terms in the last row but one. Hence, throughout the interval, since Δ vanishes, we have for every r, s

$$\Delta_r D\Delta_s - \Delta_s D\Delta_r = 0.$$

Suppose Δ_1 to be that Wronskian of the set $(\Delta_1,\ldots, \Delta_n)$ which does not

† This stipulation is not too stringent for my purpose. For other conditions see Bôcher, *Trans. American Math. Soc.* 2 (1901), 139; Curtiss, *Math. Annalen* 65 (1908), 282.

vanish anywhere in the interval. We then can write throughout the interval

$$\frac{\Delta_1 D\Delta_r - \Delta_r D\Delta_1}{\Delta_1^2} = 0 \qquad (r = 2, 3, \dots, n),$$

i.e.

$$\Delta_r/\Delta_1 = c_r, \quad \text{a constant.}$$

Again, multiplication of the terms of the first row by the minors of the last row gives

$$\Delta_1 f_1 + \Delta_2 f_2 + \dots + \Delta_n f_n = 0,$$

i.e.

$$f_1 + c_2 f_2 + \dots + c_n f_n = 0,$$

which is a linear relation connecting $f_1, f_2 \dots, f_n$.

If one of the Wronskians $\Delta_1, \dots \Delta_n$, say Δ_1, vanishes *throughout* the interval, we can start with this as the fundamental Wronskian and look for a linear relation between the $n-1$ functions $f_2, \dots, f_n$ which it involves. Such a relation would, of course, be the particular case of

$$c_1 f_1 + \dots + c_n f_n = 0$$

in which $c_1 = 0$.

One of the principal applications of the Wronskian theorem (19) is to the theory of the ordinary linear differential equation

$$X_0 D^n y + X_1 D^{n-1} y + \dots + X_n y = 0.$$

We use it to show that the general solution of such a differential equation is expressible in terms of n independent particular solutions—or, more precisely, is expressible *linearly* in terms of n *linearly* independent solutions. Points of discontinuity (and, in particular, infinities) of the coefficients $X_0, \dots, X_n$ should naturally be excluded from the domain in which we seek a solution. By considering the differential equation in the form

$$D^n y + \frac{X_1}{X_0} D^{n-1} y + \dots + \frac{X_n}{X_0} y = 0,$$

we see that zeros of X_0 should also be excluded. In the actual statement of the theorem it is, as a matter of fact, sufficient to exclude zeros of X_0 and infinities of X_1. Thus:

(20) *If $\eta_1, \dots, \eta_n$ are n linearly independent solutions of the differential equation*

$$X_0 D^n y + X_1 D^{n-1} y + \dots + X_n y = 0,$$

then, throughout an interval containing neither zeros of X_0 nor infinities of X_1, the general solution of the differential equation is given by

$$y = c_1 \eta_1 + \dots + c_n \eta_n,$$

where $c_1, \dots, c_n$ are arbitrary constants.

4. Wronskians as algebraic eliminants

The Wronskian form has useful applications to the theory of algebraic elimination. Given two polynomials $f(x)$, $g(x)$ of respective degrees m, n, we form the Wronskian of the $m+n$ functions

$$f(x),\quad xf(x),\ \ldots,\quad x^{n-1}f(x),\quad g(x),\quad xg(x),\ \ldots,\quad x^{m-1}g(x). \tag{21}$$

It is arithmetically convenient to divide each row of rth derivatives by the factor $r!$ and also to rearrange rows and columns. We then consider the Wronskian in the form

$$W[f,g] \equiv \{D,1\}^{m+n-1}|(x,1)^{n-1}f(x),\ (x,1)^{m-1}g(x)|, \tag{22}$$

where $\{D,1\}^{m+n-1}$ means $\dfrac{D^{m+n-1}}{(m+n-1)!}, \dfrac{D^{m+n-2}}{(m+n-2)!}, \ldots, 1$, and, of course, $(x,1)^{n-1}$ means $x^{n-1}, x^{n-2}, \ldots, 1$, etc. To differentiate this Wronskian we differentiate the row represented by D^{m+n-1}. But this is a row of mere constants (mostly zero), since the original functions (21) are of degree $m+n-1$ at most. Thus the derivative of the Wronskian vanishes, and the Wronskian itself is therefore independent of x.

By a similar argument the minors of the last row

$$\{D^{m+n-1},\ldots,D\}||(x,1)^{n-1}f(x),\ (x,1)^{m-1}g(x)||$$

are also constants. If we write them $A_0,\ldots,A_{n-1}, B_0,\ldots,B_{m-1}$ and expand the Wronskian by its last row, we have the formula

$$(A_0x^{n-1}+\ldots+A_{n-1})f(x)+(B_0x^{m-1}+\ldots+B_{m-1})g(x) = W, \text{ a constant.} \tag{23}$$

This is a formula well-known in the theory of algebraic eliminants, and is closely connected with the expression of $f(x)/g(x)$ as an algebraic continued fraction.

If $f(x)$, $g(x)$ vanish simultaneously for some $x=\xi$, so also does the last row of the Wronskian and therefore also the Wronskian itself. Since the Wronskian is independent of x, its vanishing is a necessary condition for the existence of a common zero of $f(x)$, $g(x)$, as is otherwise clear from (23). Thus $W[f,g]$ contains the eliminant of $f(x)$, $g(x)$ as a factor. Its dimensions in the coefficients of $f(x)$, $g(x)$ are those of the eliminant, namely n, m respectively. It is thus the eliminant itself save possibly for a numerical factor. Actually we may identify it with the eliminant

$$\begin{vmatrix} a_0 & 0 & . & b_0 & 0 & . \\ a_1 & a_0 & . & b_1 & b_0 & . \\ . & . & . & . & . & . \\ a_m & a_{m-1} & . & b_m & b_{m-1} & . \\ 0 & a_m & . & b_{m+1} & b_m & . \\ . & . & . & . & . & . \end{vmatrix} \quad (m<n), \tag{24}$$

where $f(x) \equiv a_0 x^m + a_1 x^{m-1} + \ldots + a_m$, $g(x) \equiv b_0 x^n + b_1 x^{n-1} + \ldots + b_n$. For, since W is independent of x, we may put $x = 0$ in W. A typical element of W is now

$$\left[\frac{D^r}{r!}\{x^s f(x)\}\right]_{x=0}.$$

But this is exactly the coefficient of x^r in $x^s f(x)$, since terms of order less than r disappear under the differentiation and terms of order greater than r disappear when x is put equal to zero. The term in question is thus just the coefficient of x^{r-s} in $f(x)$. If every element of W is now replaced in this way, it will be found that we have precisely the eliminant in the form (24).

We go on to consider more generally the series of determinants

$$R_{s-1}(x) \equiv \{D^{m+n-s}, D^{m+n-s-1},\ldots, D^s, 1\}|(x,1)^{n-s}f(x),\ (x,1)^{m-s}g(x)|$$

where $s = 1,\ldots,m$, if $m \leqslant n$. The first of them $R_0(x)$ is, of course, the Wronskian $W[f,g]$ itself. The others are not Wronskians in the strict sense, but it is evident that we differentiate $R_{s-1}(x)$ by differentiating its first and last rows in turn. The first row, symbolized by D^{m+n-s}, is a row of constants (mostly zero), and so we have

$$DR_{s-1}(x) = \{D^{m+n-s}, D^{m+n-s-1},\ldots, D^s, D\}|(x,1)^{n-s}f(x),\ (x,1)^{m-s}g(x)|$$

and generally

$$D^r R_{s-1}(x) = \{D^{m+n-s}, D^{m+n-s-1},\ldots, D^s, D^r\}|(x,1)^{n-s}f(x), (x,1)^{m-s}g(x)|.$$

Hence $D^s R_{s-1}(x) = 0$ and therefore $R_{s-1}(x)$ is a polynomial of degree not exceeding $s-1$.

The minors of the last row in $R_{s-1}(x)$ are

$$\{D^{m+n-s}, D^{m+n-s-1},\ldots, D^s\}||(x,1)^{n-s}f(x),\ (x,1)^{m-s}g(x)||$$

which, by an argument already used, are all constants. Calling them (in order)

$$_sA_0,\quad {}_sA_1,\quad \ldots,\quad {}_sA_{n-s},\quad {}_sB_0,\quad {}_sB_1,\quad \ldots,\quad {}_sB_{m-s},$$

define

$$(-)^{m+n}f_s(x) \equiv {}_sB_0 x^{m-s} + {}_sB_1 x^{m-s-1} + \ldots + {}_sB_{m-s},$$

$$(-)^{m+n-1}g_s(x) \equiv {}_sA_0 x^{n-s} + {}_sA_1 x^{n-s-1} + \ldots + {}_sA_{n-s},$$

so that, slightly extending our notation, we may write

$$-f_s(x) = \{D^{m+n-s}, D^{m+n-s-1},\ldots, D^s\}\begin{vmatrix} (x,1)^{n-s}f(x), & (x,1)^{m-s}g(x) \\ (x,1)^{n-s}, & 0,\ldots,0 \end{vmatrix}, \qquad (25)$$

$$g_s(x) = \{D^{m+n-s}, D^{m+n-s-1},\ldots, D^s\}\begin{vmatrix} (x,1)^{n-s}f(x), & (x,1)^{m-s}g(x) \\ 0,\ldots,0, & (x,1)^{m-s} \end{vmatrix}.$$

Thus expansion of $R_{s-1}(x)$ by its last row gives

$$R_{s-1}(x) = f(x)g_s(x) - g(x)f_s(x), \tag{26}$$

where $R_{s-1}(x)$, $f_s(x)$, $g_s(x)$ are polynomials of degrees not exceeding $s-1$, $m-s$, $n-s$, respectively.

It is fundamental in the algebraic theory that with suitable constant multipliers, $R_{m-1}(x)$, $R_{m-2}(x)$,..., $R_0(x)$ are the successive remainders in the conversion of $g(x)/f(x)$ into a continued fraction, while

$$g_m(x)/f_m(x), \quad g_{m-1}(x)/f_{m-1}(x), \quad \ldots, \quad g_1(x)/f_1(x)$$

are the successive convergents to the continued fraction.

Consequences of importance in the theory of elimination follow from (25): for instance, if $f(x)$, $g(x)$ have a common factor of order k, then $R_0(x)$, $R_1(x)$,..., $R_{k-1}(x)$ must all vanish identically and the factor itself is sufficiently given by $R_k(x)$.

The algebraic theory is given at length by Sylvester.†

5. Generalized theorem of the mean. Taylor's series

The theorem of mean value proved in chapter IV § 11 (48) for the first derivative generalizes for the nth derivative into the form

(27) *If $f_n(x)$, $g_n(x)$ exist throughout $(a, a+h)$ and have no common zero therein, then*

$$\frac{f(a+h) - \sum_{r=0}^{n-1} \frac{h^r}{r!} f_r(a)}{g(a+h) - \sum_{r=0}^{n-1} \frac{h^r}{r!} g_r(a)} = \frac{f_n(a+\theta h)}{g_n(a+\theta h)},$$

for some θ in $)0, 1($.

In this formula we have conveniently written $f_n(x)$, $g_n(x)$ for the nth derivatives of $f(x)$, $g(x)$ and in particular $f_0(x)$ $g_0(x)$ for the functions $f(x)$, $g(x)$ themselves. To prove (27) write

$$\phi(t) \equiv \sum_{r=0}^{n-1} \frac{(h-t)^r}{r!} f_r(a+t), \qquad \psi(t) \equiv \sum_{r=0}^{n-1} \frac{(h-t)^r}{r!} g_r(a+t).$$

Then
$$\begin{aligned}\phi'(t) &= \sum_{r=0}^{n-1} \frac{(h-t)^r}{r!} f_{r+1}(a+t) - \sum_{r=1}^{n-1} \frac{(h-t)^{r-1}}{(r-1)!} f_r(a+t) \\ &= \sum_{r=0}^{n-1} \frac{(h-t)^r}{r!} f_{r+1}(a+t) - \sum_{r=0}^{n-2} \frac{(h-t)^r}{r!} f_{r+1}(a+t) \\ &= \frac{(h-t)^{n-1}}{(n-1)!} f_n(a+t).\end{aligned}$$

† *Phil. Trans.* 143 (1853), 407–548 or *Collected Works*, 1, 427 et seq.

But $$\phi(h) = f(a+h), \qquad \phi(0) = \sum_{r=0}^{n-1} \frac{h^r}{r!} f_r(a).$$

Similarly, $$\psi(h) = g(a+h), \qquad \psi(0) = \sum_{r=0}^{n-1} \frac{h^r}{r!} g_r(a),$$

$$\psi'(t) = \frac{(h-t)^{n-1}}{(n-1)!} g_n(a+t).$$

Thus $$\frac{f(a+h) - \sum\limits_{r=0}^{n-1} \frac{h^r}{r!} f_r(a)}{g(a+h) - \sum\limits_{r=0}^{n-1} \frac{h^r}{r!} g_r(a)} = \frac{\phi(h)-\phi(0)}{\psi(h)-\psi(0)}$$

$$= \frac{\phi'(\theta h)}{\psi'(\theta h)}, \quad \text{by IV (48)},$$

$$= \frac{f_n(a+\theta h)}{g_n(a+\theta h)}, \quad \text{where} \quad 0 < \theta < 1.$$

We must stipulate that $\phi'(t)$, $\psi'(t)$ have no common zero in $(0, h)$, i.e. that $f_n(x)$, $g_n(x)$ have no common zero in $(a, a+h)$. This is the generalized *fractional* formula of the mean.

To get the corresponding linear formula we write

$$g(x) = (x-a)^n,$$

so that $$g_r(a) = 0 \quad (r = 0, 1, \ldots, n-1), \qquad g_n(a) = n!$$

We thus have the theorem:

(28) *If $f_n(x)$ exists throughout $(a, a+h)$, then*

$$f(a+h) - \sum_{r=0}^{n-1} \frac{h^r}{r!} f_r(a) = \frac{h^n f_n(a+\theta h)}{n!},$$

for some θ in $)0, 1($.

The stipulation that $f_n(x)$, $g_n(x)$ have no common zero in $(a, a+h)$ is now satisfied automatically, since $g_n(a)$ is a constant other than zero.

The series $\sum h^r f_r(a)/r!$ associated with a given function $f(x)$ is known as a *Taylor's series* of the function. In certain circumstances, which we shall discuss in a later chapter, it can give a representation of $f(a+h)$. It is clear from (28) that, if h is small, the first n terms of the series give an approximation to $f(a+h)$ correct to h^n.

From this point of view the formula (28) is a formula for *the remainder after n terms* in Taylor's series. Write $R_n(a,h)$ for this remainder, i.e.

$$R_n(a,h) \equiv f(a+h) - \sum_{r=0}^{n-1} \frac{h^r}{r!} f_r(a).$$

Thus (28) gives
$$R_n(a,h) = \frac{h^n}{n!} f_n(a+\theta h), \tag{29}$$
a form due to Lagrange.

Again, in the proof of (27), had we written $\psi(t) \equiv t$, i.e. had we applied the linear formula of the mean† to $\phi(t)$, we should have obtained

$$R_n(a,h) = \phi(h) - \phi(0) = h\phi'(\theta h),$$

that is to say
$$R_n(a,h) = \frac{(1-\theta)^{n-1} h^n f_n(a+\theta h)}{(n-1)!}, \tag{30}$$
a form due to Cauchy.

Evidently, with $\psi(t)$ at our disposal, we have endless variety of possible forms of R_n, so long as we observe the single condition that $\psi'(t)$ do not vanish in $)0,h($. For otherwise we should have to impose on $f(x)$ the condition that $f_n(x)$ do not vanish in $(a, a+h)$.

All such expressions are essentially approximations to R_n, since their virtue lies in the expressed or implied inequality $0 < \theta < 1$. It should be remembered that the Integral Calculus enables us to state a simple and exact form for $R_n(a,h)$.

We may use (28) to give an immediate proof of (7), namely that, in an interval throughout which $f_n(x)$ vanishes, $f(x)$ is a polynomial of degree less than n. For, in (28), the expression on the right now vanishes throughout the interval and so exactly

$$f(a+h) = \sum_{r=0}^{n-1} \frac{h^r}{r!} f_r(a)$$

which gives $f(a+h)$ as a polynomial in h of degree less than n.

6. Taylor's limit

It follows at once from (27) and (28) that, if $f_n(x)$, $g_n(x)$ are continuous at $x = a$ then

$$\lim_{h\to 0} \frac{f(a+h) - \sum_{r=0}^{n-1} \frac{h^r}{r!} f_r(a)}{g(a+h) - \sum_{r=0}^{n-1} \frac{h^r}{r!} g_r(a)} = \frac{f_n(a)}{g_n(a)}, \tag{31}$$

† Chapter IV (49).

and, in particular,
$$\lim_{h\to 0}\frac{f(a+h)-\sum\limits_{r=0}^{n-1}\frac{h^r}{r!}f_r(a)}{h^n}=\frac{f_n(a)}{n!}. \tag{32}$$

But it is not necessary to the existence of these limits that $f_n(x)$, $g_n(x)$ be continuous at $x = a$ or even exist near $x = a$. It is sufficient merely that $f_n(a)$, $g_n(a)$ themselves exist. For write
$$F(x) \equiv f(x)-\frac{(x-a)^{n-1}}{(n-1)!}f_{n-1}(a).$$

Then
$$F_r(a) = f_r(a) \qquad (r = 0, 1,\ldots, n-2)$$
and
$$F_{n-1}(x) = f_{n-1}(x)-f_{n-1}(a).$$

Thus
$$f(a+h)-\sum_{r=0}^{n-1}\frac{h^r}{r!}f_r(a) = F(a+h)-\sum_{r=0}^{n-2}\frac{h^r}{r!}F_r(a).$$

With $G(x)$ similarly defined in terms of $g(x)$ we have that
$$\frac{f(a+h)-\sum\limits_{r=0}^{n-1}\frac{h^r}{r!}f_r(a)}{g(a+h)-\sum\limits_{r=0}^{n-1}\frac{h^r}{r!}g_r(a)}=\frac{F(a+h)-\sum\limits_{r=0}^{n-2}\frac{h^r}{r!}F_r(a)}{G(a+h)-\sum\limits_{r=0}^{n-2}\frac{h^r}{r!}G_r(a)}$$
$$=\frac{F_{n-1}(a+\theta h)}{G_{n-1}(a+\theta h)}, \text{ by (27) with } n-1 \text{ for } n,$$
$$=\frac{f_{n-1}(a+\theta h)-f_{n-1}(a)}{g_{n-1}(a+\theta h)-g_{n-1}(a)}, \text{ on substitution for } F_{n-1}, G_{n-1},$$
$$=\frac{\{f_{n-1}(a+\theta h)-f_{n-1}(a)\}/\theta h}{\{g_{n-1}(a+\theta h)-g_{n-1}(a)\}/\theta h}$$
$$\to\frac{f_n(a)}{g_n(a)} \text{ as } h\to 0, \text{ since the } n\text{th derivatives exist.}$$

As always in the fractional form we must exclude the possibility that $F_{n-1}(x)$, $G_{n-1}(x)$ vanish together in $(a, a+h)$. Since we are concerned only with the limit $h \to 0$, we can choose h small enough to exclude such a common zero, unless $F_{n-1}(x)$, $G_{n-1}(x)$ have common zeros in every neighbourhood of $x = a$. In such a case
$$\frac{f_{n-1}(x)-f_{n-1}(a)}{x-a}, \qquad \frac{g_{n-1}(x)-g_{n-1}(a)}{x-a}$$
both vanish in every neighbourhood of $x = a$ and therefore their limits,

as $x \to a$, also vanish. In this exceptional case the stated limit $f_n(a)/g_n(a)$ assumes the meaningless form 0/0 and should therefore in any case be excluded.

If one only of $f_n(a)$, $g_n(a)$ vanishes, we must similarly secure that it appears in the numerator and not the denominator of the stated limit. We may therefore enunciate the theorem in the form:

(33) *If $f_n(a)$, $g_n(a)$ exist and if $g_n(a) \neq 0$,* then

$$\lim_{h\to 0}\frac{f(a+h)-\sum_{r=0}^{n-1}\frac{h^r}{r!}f_r(a)}{g(a+h)-\sum_{r=0}^{n-1}\frac{h^r}{r!}g_r(a)}=\frac{f_n(a)}{g_n(a)}.$$

In particular, if we write

$$g(x)\equiv(x-a)^n,$$

so that

$$g_r(a)=0 \qquad (r=1,\dots,n-1),$$

$$g_n(a)=n!\neq 0,$$

we have similarly

$$\lim_{h\to 0}h^{-n}\left\{f(a+h)-\sum_{r=0}^{n-1}\frac{h^r}{r!}f_r(a)\right\}=\frac{f_n(a)}{n!}, \tag{34}$$

provided only that $f_n(a)$ exists.

This result may be called the *Taylor's limit of order n*. It is often convenient to write it in the equivalent form

$$f(a+h)=\sum_{r=0}^{n}\frac{h^r}{r!}f_r(a)+h^n\epsilon(a,h), \tag{35}$$

where $\epsilon(a,h)\to 0$ as $h\to 0$.

7. Deductions from Taylor's limit

We may deduce from (35) a formula of rather different character:

$$f(a)=\sum_{r=0}^{n-1}\frac{(a-x)^rf_r(x)}{r!}+\frac{(a-x)^nf_n(a)}{n!}+(a-x)^n\epsilon(a,x), \tag{36}$$

where $\epsilon(a,x)\to 0$ as $x\to 0$, provided only that $f_n(a)$ exists.

For using (35) we may write

$$\sum_{r=0}^{n-1}\frac{(a-x)^rf_r(x)}{r!}+\frac{(a-x)^nf_n(a)}{n!}$$

$$=\sum_{r=0}^{n-1}\left\{\frac{(a-x)^r}{r!}\left[\sum_{s=0}^{n-r}\frac{(x-a)^sf_{r+s}(a)}{s!}+(x-a)^{n-r}\epsilon_r(a,x)\right]\right\}+\frac{(a-x)^nf_n(a)}{n!},$$

where $\epsilon_0(a,x),\ldots,\epsilon_{n-1}(a,x) \to 0$ as $x \to 0$. By rearranging the repeated summation we may write this as

$$f(a)+\sum_{p=1}^{n}(x-a)^p f_p(a)\sum_{q=0}^{p}\frac{(-1)^q}{q!(p-q)!}-(a-x)^n\epsilon(a,x),$$

where

$$\epsilon(a,x) \equiv (-)^{n-1}\sum_{r=0}^{n-1}\frac{\epsilon_r(a,x)}{r!}$$

and accordingly $\epsilon(a,x) \to 0$ as $x \to 0$. Since

$$\sum_{q=0}^{p}\frac{(-1)^q}{q!(p-q)!}=\frac{(1-1)^p}{p!}=0,$$

we have the required formula.

From (36) we deduce a result which we shall need subsequently:

(37) *If $f(x),\ldots,k(x)$ are n functions which vanish at $x = a$ and are n times differentiable there, then*

$$\lim_{x\to a}\frac{(1,D,\ldots,D^{n-1})|f(x)\ldots k(x)|}{(x-a)^n}=\frac{(D,D^2,\ldots,D^n)|f(a)\ldots k(a)|}{n!}.$$

For, if in the determinant

$$(1,D,\ldots,D^{n-1})|f(x)\ldots k(x)|$$

we multiply the second, third,..., nth rows by

$$\frac{a-x}{1!},\quad \frac{(a-x)^2}{2!},\ldots,\quad \frac{(a-x)^{n-1}}{(n-1)!}$$

respectively and add to the first row, we get

$$\sum_{r=0}^{n-1}\frac{(a-x)^r f_r(x)}{r!},\ldots,\quad \sum_{r=0}^{n-1}\frac{(a-x)^r k_r(x)}{r!},$$

which, since $f(a),\ldots,k(a)$ all vanish, may be written, in virtue of (36), as

$$-(a-x)^n\left\{\frac{f_n(a)}{n!}+\epsilon(a,x)\right\},\ldots,\quad -(a-x)^n\left\{\frac{k_n(a)}{n!}+\omega(a,x)\right\},$$

where $\epsilon,\ldots,\omega \to 0$ as $x \to a$.

In $$(1,D,\ldots,D^{n-1})|f(x)\ldots k(x)|/(x-a)^n$$

we may therefore divide by $(x-a)^n$ and proceed at once to the limit, which gives

$$(-)^{n-1}(D^n,D,\ldots,D^{n-1})|f(a)\ldots k(a)|/n! = (D,D^2,\ldots,D^n)|f(a)\ldots k(a)|/n!,$$

and the result is established.

We may also use (35) to give an alternative proof of Leibniz's theorem. We must first observe that, by chapter IV (15), the product $f(x)g(x)$ is differentiable, if $f(x)$, $g(x)$ are themselves differentiable, the derivative being $f'(x)g(x)+f(x)g'(x)$. Hence this derivative is also differentiable, if $f'(x)$, $g'(x)$ are also differentiable. Thus, inductively, $f(x)g(x)$ is n times differentiable, if $f(x)$, $g(x)$ are themselves n times differentiable.

Applying (35) to the product $f(u+h)g(u+h)$ and also to the separate factors we have

$$\sum_{r=0}^{n} \frac{h^r}{r!} D^r\{f(u)g(u)\} + h^n \epsilon(u,h)$$
$$= \left[\sum_{r=0}^{n} \frac{h^r}{r!} f_r(u) + h^n \epsilon_1(u,h)\right]\left[\sum_{r=0}^{n} \frac{h^r}{r!} g_r(u) + h^n \epsilon_2(u,h)\right],$$

where $\epsilon(u,h)$, $\epsilon_1(u,h)$, $\epsilon_2(u,h)$ all converge to zero, as $h \to 0$. Dividing in turn by $1, h,..., h^n$ and taking the limit $h \to 0$, we have, for $r = 0, 1,..., n$,

$$D^r\{f(u)g(u)\} = r! \sum_{s=0}^{r} \frac{f_s(u)g_{r-s}(u)}{s!(r-s)!},$$

which is Leibniz's formula.

More generally we may obtain in like fashion a formula for the nth derivative of a function of a function:

(38) *If, for corresponding values of x and $u = u(x)$, $D^n u, f_n(u)$ both exist, then $D^n f(u)$ also exists and is the coefficient of h^n in*

$$n! \sum_{r=0}^{n} \left(h\, Du + \frac{h^2}{2!} D^2 u + ... + \frac{h^n}{n!} D^n u\right)^r \frac{f_r(u)}{r!}.$$

As in the foregoing proof of Leibniz's formula, successive differentiation shows that the existence of $D^n u$, $f_n(u)$ is sufficient to imply the existence of $D^n f(u)$. Now write

$$U \equiv u(x+h), \qquad H \equiv \sum_{r=1}^{n} \frac{h^r}{r!} D^r u,$$

where the parameter h is independent of x. Then by the theorem of the mean, chapter IV (48),

$$\lim_{h\to 0} \frac{f(U)-f(u+H)}{U-(u+H)} = f'(u),$$

for $f'(u)$ is continuous, since $f''(u)$ exists, if we suppose that $n \geqslant 2$. But, by Taylor's limit (34) above,

$$\lim_{h\to 0} \frac{U-(u+H)}{h^n} = 0.$$

Hence
$$\lim_{h\to 0}\frac{f(U)-f(u+H)}{h^n} = 0.$$

But, again by Taylor's limit,
$$\lim_{h\to 0}\frac{f(U)-\sum_{r=0}^{n}\frac{h^r}{r!}D^r f(u)}{h^n} = 0$$

and
$$\lim_{h\to 0}\frac{f(u+H)-\sum_{r=0}^{n}\frac{H^r}{r!}f_r(u)}{h^n} = 0.$$

Thus
$$\lim_{h\to 0}\frac{\sum_{r=0}^{n}\frac{h^r}{r!}D^r f(u)-\sum_{r=0}^{n}\frac{H^r}{r!}f_r(u)}{h^n} = 0.$$

The numerator is a polynomial in h of degree n^2. The existence of the limit requires the absence from this numerator of terms in h whose degree does not exceed n. In particular the coefficient of h^n in the numerator must vanish, i.e. we must have

$$D^n f(u) = \text{coefficient of } h^n \text{ in } n! \sum_{r=0}^{n}\left(h\,Du+\frac{h^2}{2!}D^2u+\ldots+\frac{h^n}{n!}D^n u\right)^r\frac{f_r(u)}{r!},$$

which establishes the theorem.

8. 'Umbral' derivatives

As we have just seen, the existence of $f_n(a)$ is sufficient to secure the existence of the Taylor's limit

$$\lim_{h\to 0} h^{-n}\left\{f(a+h)-\sum_{r=0}^{n-1}\frac{h^r}{r!}f_r(a)\right\}.$$

But conversely the existence of $f_n(a)$ is not necessary to the existence of the Taylor's limit. For example, taking $n = 2$, write

$$\left.\begin{aligned} f(x) &= x^3\sin(x^{-2}) \qquad (x\neq 0)\\ f(0) &= 0.\end{aligned}\right\} \tag{39}$$

Then, except at the origin,

$$f'(x) = 3x^2\sin(x^{-2})-2\cos(x^{-2}),$$

which is discontinuous at $x = 0$ and so $f''(0)$ certainly does not exist.

But
$$f'(0) = \lim_{h\to 0}\frac{h^3\sin(h^{-2})}{h} = 0.$$

Hence
$$\frac{f(h)-f(0)-hf'(0)}{h^2} = h\sin(h^{-2}),$$

and this tends to 0 as $h \to 0$. Thus the Taylor's limit of the second order exists at $x = 0$, although $f''(0)$ does not exist.

We may go farther and construct a function in which $f'(x)$ exists only at $x = 0$ and yet the Taylor's limit of the second order also exists at $x = 0$. Write

$$\left.\begin{aligned} f(x) &\equiv 0 && (x \text{ irrational or zero}) \\ f(x) &\equiv p/q^2\sqrt{q} && (x \text{ rational, in lowest terms } p/q \text{ where } q > 0). \end{aligned}\right\} \quad (40)$$

Then, if h is irrational,

$$\frac{f(h)-f(0)}{h} = 0;$$

if $h = p/q$, $$\frac{f(h)-f(0)}{h} = \frac{1}{q\sqrt{q}} < h \text{ numerically.}$$

Hence $f'(0) = 0$, and

$$h^{-2}\{f(h)-f(0)-hf'(0)\} = h^{-2}f(h) = 0, \text{ if } h \text{ is irrational;}$$

and $$= 1/p\sqrt{q}, \text{ if } h = p/q < \sqrt{|h|}.$$

Hence $$h^{-2}\{f(h)-f(0)-hf'(0)\} \to 0 \text{ as } h \to 0.$$

But $f(x)$ is nowhere differentiable except at the origin. At any other rational point p/q it is discontinuous (like all of its type), since if we approach p/q by a sequence of always irrational arguments, $f(x)$ is always zero and so does not converge to $p/q^2\sqrt{q}$.

If ξ is irrational, suppose it expressed as a simple continued fraction and let $\xi+h$ approach ξ through the sequence of convergents p/q of this continued fraction. If p'/q' denote the convergent immediately preceding p/q and α the complete quotient immediately following, we have, by the usual theory,†

$$h = \frac{p}{q} - \xi = \pm\frac{1}{q(\alpha q+q')}.$$

Hence $$\left|\frac{f(\xi+h)-f(\xi)}{h}\right| = \left|\frac{p(\alpha q+q')}{q\sqrt{q}}\right|$$

$$> |\xi+h|\sqrt{q}, \text{ since } \alpha > 1$$

Thus the incremental ratio diverges for this sequence of approach. If $\xi+h$ approach ξ through values always irrational, the incremental ratio is always zero. Evidently the derivative exists for no irrational argument.

Hence, in sum, the first derivative nowhere exists except at $x = 0$, and yet the Taylor's limit of the second order exists at $x = 0$. We may

† Cf. Chrystal, *Algebra*, 2 (1900), 437.

regard this limit as giving a sort of simulacrum of the second-order derivative at the point. Assuming the limit to exist I shall write

$$\lim_{h\to 0}\frac{f(a+h)-f(a)-hf'(a)}{h^2/2!}=[f''(a)],$$

and I shall call $[f''(a)]$ the *umbral derivative*† of the second order at $x = a$. We may similarly define the umbral derivative of the third order as

$$\lim_{h\to 0}\frac{f(a+h)-f(a)-h\,f'(a)-\dfrac{h^2}{2!}[f''(a)]}{h^3/3!}=[f'''(a)],$$

assuming the limit to exist.

Proceeding inductively in this way, we at length define the umbral derivative of the nth order at $x = a$ by the inductive formula

$$\lim_{h\to 0}\frac{f(a+h)-f(a)-h\,f'(a)-\sum\limits_{r=2}^{n-1}\dfrac{h^r}{r!}[f_r(a)]}{h^n/n!}=[f_n(a)]. \tag{41}$$

Of course, if $f(x)$ is differentiable $n-m$ times (in the ordinary sense) at $x = a$, we may replace the umbral derivatives of orders $2, 3,\ldots, n-m$ by their proper derivatives.

As an example of an nth umbral derivative, consider the function

$$\left.\begin{aligned} &f(x)=\exp(-x^{-2})\sin\{\exp(x^{-2})\} \qquad (x\neq 0)\\ &f(0)=0\end{aligned}\right\}. \tag{42}$$

We have $f'(0) = 0$ and, if $x \neq 0$,

$$f'(x)=2x^{-3}\exp(-x^{-2})\sin\{\exp(x^{-2})\}-2x^{-3}\cos\{\exp(x^{-2})\}.$$

Hence, at the origin, $f'(x)$ oscillates infinitely and $f''(0)$ does not exist.

But $$[f''(0)]=2\lim_{x\to 0}\frac{f(x)-f(0)-xf'(0)}{x^2}=0,$$

and, proceeding inductively, we find that, for every n,

$$[f_n(0)]=n!\lim_{x\to 0}\frac{f(x)}{x^n}=0.$$

We may extend the result of (38) to umbral derivatives as follows:

(43) *If for corresponding values of x and $u = u(x)$, the umbral derivatives $[D^n u]$, $[f_n(u)]$ exist, then the umbral derivative $[D^n f(u)]$ also exists and is equal to the coefficient of h^n in*

$$n!\sum_{r=0}^{n}\left\{h\,Du+\frac{h^2}{2!}[D^2u]+\ldots+\frac{h^n}{n!}[D^n u]\right\}^r\frac{[f_r(u)]}{r!}.$$

† The point (if any) of the umbral derivative is that it serves to remind us how many properties of the nth derivative arise simply from its being a coefficient in Taylor's series. This we shall see subsequently.

For, if $U \equiv u(x+h)$, then, since $[f_n(u)]$ exists,

$$f(U) = \sum_{r=0}^{n-1} \frac{(U-u)^r}{r!}[f_r(u)] + \frac{(U-u)^n}{n!}\{[f_n(u)]+\eta\},$$

where $\eta \to 0$ as $U \to u$. Again, since $[D^n u]$ exists,

$$U = \sum_{r=0}^{n-1} \frac{h^r}{r!}[D^r u] + \frac{h^n}{n!}\{[D^n u]+\epsilon\},$$

where $\epsilon \to 0$ as $h \to 0$. Hence

$$f(U) = \sum_{r=0}^{n-1} \left\{h\, Du + \dots + \frac{h^{n-1}}{(n-1)!}[D^{n-1}u] + \frac{h^n}{n!}([D^n u]+\epsilon)\right\}^r \frac{[f_r(u)]}{r!} + $$
$$+ \left\{h\, Du + \dots + \frac{h^{n-1}}{(n-1)!}[D^{n-1}u] + \frac{h^n}{n!}([D^n u]+\epsilon)\right\}^n \frac{[f_n(u)]+\eta}{n!}. \quad (44)$$

Adopting an inductive method of proof we suppose the theorem to hold for values of the index up to $n-1$ inclusive. Then

$$\sum_{r=0}^{n-1} \frac{h^r}{r!} D^r[f(u)]$$

consists of the terms up to h^{n-1} inclusive in the expansion in powers of h of

$$\sum_{r=0}^{n-1} \left\{h\, Du + \dots + \frac{h^{n-1}}{(n-1)!}[D^{n-1}u]\right\}^r \frac{[f_r(u)]}{r!}.$$

These are equally the terms up to h^{n-1} in the similar expansion for $f(U)$ in (44). Hence

$$\frac{f(U) - \sum_{r=0}^{n-1} \frac{h^r}{r!} D^r[f(u)]}{h^n/n!} \quad (45)$$

contains no negative powers of h. The only other functions of h which it involves are ϵ, η, which both converge (to zero) as $h \to 0$. Thus (45) converges as $h \to 0$. Its limit is, by definition, the umbral derivative $[D^n f(u)]$, which therefore exists, if the theorem (43) holds for values of the index up to $n-1$. To evaluate the limit it is evidently enough to retain only the coefficient of h^n in the numerator of (45). Thus

$$[D^n f(u)] = n! \lim_{h \to 0} \{\text{coefficient of } h^n \text{ in } f(U)\}$$
$$= n! \times \text{coefficient of } h^n \text{ in } \sum_{r=0}^{n} \left\{h\, Du + \dots + \frac{h^n}{n!}[D^n u]\right\}^r \frac{f_r(u)}{r!},$$

when we take the limits $\epsilon, \eta \to 0$.

Hence, if (43) is true for indices up to $n-1$, it is also true for the index n. It is evidently true for $n = 1$, and is therefore true generally.

9. An incremental ratio for the second derivative

Our discussion of the derivative of $f(x)$ has been based throughout on its definition as the limit of the incremental ratio

$$\frac{f(x+h)-f(x)}{h},$$

and in this chapter the higher derivatives have been defined successively as limits of incremental ratios of their predecessors, i.e. by the formula

$$f_{n+1}(x) = \lim_{h\to 0}\frac{f_n(x+h)-f_n(x)}{h},$$

which is essentially a recurrence-formula. It is, however, possible to construct an appropriate incremental ratio from which we can proceed directly to the nth derivative. Beginning with the second derivative as sufficiently typical I write for the basic incremental ratio

$$\phi(x,h) = \frac{f(x+h)-f(x)}{h}. \tag{46}$$

Regarding $\phi(x,h)$ as a function of x only, form its incremental ratio with increment k and write

$$\phi(x,h,k) = \frac{f(x+h+k)-f(x+h)-f(x+k)+f(x)}{hk}. \tag{47}$$

We can prove that

(48) *If both the limits*

$$\lim_{h\to 0}\phi(\xi,h), \quad \lim_{h\to 0}\lim_{k\to 0}\phi(\xi,h,k)$$

exist, then $f''(\xi)$ exists and is equal to the repeated limit. Conversely, if $f''(\xi)$ exists, then both the given limits exist and the repeated limit is equal to $f''(\xi)$.

The order of taking the limits $h \to 0$, $k \to 0$ is, of course, immaterial, since $\phi(\xi,h,k)$ is symmetrical in h, k.

Now $$\phi(\xi,h,k) = \frac{\phi(\xi+k,h)-\phi(\xi,h)}{k},$$

and so $$\lim_{k\to 0}\phi(\xi,h,k) = \left\{\frac{d}{dx}\phi(x,h)\right\}_{x=\xi}.$$

Thus $\phi(x,h)$, i.e. $\dfrac{f(x+h)-f(x)}{h}$, is differentiable at $x = \xi$.

If also $\lim\limits_{h\to 0}\phi(\xi,h)$ exists, $f(x)$ is differentiable at $x=\xi$, and therefore $f(x+h)$ is also differentiable at $x=\xi$. Hence

$$\lim_{k\to 0}\phi(\xi,h,k)=\frac{f'(\xi+h)-f'(\xi)}{h}.$$

Hence further $$\lim_{h\to 0}\lim_{k\to 0}\phi(\xi,h,k)=f''(\xi).$$

This establishes the direct theorem.

Conversely, if $f''(\xi)$ exists and therefore also, by implication, $f'(x)$ exists near $x=\xi$, then

$$\lim_{k\to 0}\phi(\xi,h,k)=\frac{f'(\xi+h)-f'(\xi)}{h},$$

and so $$\lim_{h\to 0}\lim_{k\to 0}\phi(\xi,h,k)=f''(\xi).$$

It should be remarked that, in the enunciation and proof of (48), it has been tacitly assumed that the existence of the limit

$$\lim_{h\to 0}\lim_{k\to 0}\phi(\xi,h,k)$$

is not sufficient to secure the existence of the limit

$$\lim_{h\to 0}\phi(\xi,h).$$

This is equivalent to the assumption that

(α) '$f(x+h)-f(x)$ is differentiable at $x=\xi$ for every h'

does not imply

(β) '$f(x)$ is differentiable at $x=\xi$'.

I cannot deduce (β) from (α): on the other hand, I cannot construct a function for which (α) is true and (β) false. Moreover, the position is the same, if, in (α) and (β), 'differentiable' be replaced by 'continuous'.†

We may further prove that

(49) *If $f''(\xi)$ exists, then $\phi(\xi,h,k)$ is convergent in (h,k) at $(0,0)$.*

For write $$F(t)\equiv\frac{f(\xi+t)-f(\xi)}{t},$$

then
$$\phi(\xi,h,k)=\frac{F(h+k)-F(h)}{k}+\frac{F(h+k)-F(k)}{h}$$
$$=F'(h+\theta_1 k)+F'(k+\theta_2 h),$$

† Prof. Pólya has pointed out to me that a discontinuous function can be constructed to satisfy the identity
$$f(x+y)=f(x)+f(y),$$
if Zermelo's hypothesis is assumed. (Hamel, *Math. Annalen*, 60 (1905), 459.) For such a function, $f(x+h)-f(x)$ is constant and so is both continuous and differentiable, although $f(x)$ is discontinuous and therefore not differentiable.

where θ_1, θ_2 lie in $)0,1($; for $F(t)$ is differentiable, since $f(x)$ is differentiable at and near $x = \xi$.

Actually
$$F'(t) = \frac{f'(\xi+t)}{t} - \frac{f(\xi+t)-f(\xi)}{t^2}$$
$$= \frac{f'(\xi+t)-f'(\xi)}{t} - \frac{f(\xi+t)-f(\xi)-tf'(\xi)}{t^2}$$
$$\to f''(\xi)-\tfrac{1}{2}f''(\xi) \quad \text{as} \quad t \to 0.$$

Hence
$$\phi(\xi,h,k) \to f''(\xi) \quad \text{as} \quad h,k \to 0.$$

It follows from (48), (49) that

(50) *If both the limits*
$$\lim_{h\to 0} \phi(\xi,h), \quad \lim_{h\to 0} \lim_{k\to 0} \phi(\xi,h,k),$$
exist, then $\phi(\xi,h,k)$ is convergent in (h,k) at $(0,0)$. i.e. the double limit
$$\lim_{h,k\to 0} \phi(\xi,h,k)$$
exists.

Suppose now that $f''(x)$ exists not only at $x = \xi$ but also in its neighbourhood. We can then raise the question of its continuity and answer it in the theorem:

(51) *The continuity of $f''(x)$ at $x = \xi$ implies and is implied by the continuity of $\phi(x,h,k)$ in (x,h,k) at $(\xi,0,0)$.*

For
$$\phi(x,h) = f'(x+\theta_1 h), \text{ where } 0 < \theta_1 < 1,$$
and so
$$\phi(x,h,k) = \frac{f'(x+\theta_1 h+k)-f'(x+\theta_1 h)}{k}$$
$$= f''(x+\theta_1 h+\theta_2 k), \text{ where } 0 < \theta_2 < 1.$$

Hence ϕ is continuous in (x,h,k) at $(\xi,0,0)$, if $f''(x)$ is continuous at $x = \xi$.

Conversely, if ϕ is continuous, we have
$$|\phi(x,h,k)-\phi(\xi,h,k)| < \epsilon,$$
if $|x-\xi|$, $|h|$, $|k| <$ some $\eta(\epsilon,\xi)$.

Since $f''(x)$ and $f''(\xi)$ exist, we may take the repeated limit $h \to 0$, $k \to 0$ in this inequality. This gives
$$|f''(x)-f''(\xi)| \leqslant \epsilon, \quad \text{if } |x-\xi| < \eta(\epsilon,\xi),$$
and establishes the continuity of $f''(x)$ at $x = \xi$.

If we had formed the incremental ratio of
$$\phi(x,h) = \frac{f(x+h)-f(x)}{h}$$

with regard to h instead of with regard to x, the increment being $k-h$, we should have got the function

$$\psi(x,h,k) \equiv \frac{kf(x+h)-hf(x+k)+(h-k)f(x)}{hk(h-k)}. \tag{52}$$

Now we may write

$$\psi(\xi,h,k) = \frac{f(\xi+h)-f(\xi)}{h(h-k)} - \frac{f(\xi+k)-f(\xi)}{k(h-k)}.$$

Hence $\lim\limits_{k\to 0}\psi$ exists, if and only if $f'(\xi)$ exists, and then we have

$$\lim_{k\to 0}\psi(\xi,h,k) = \frac{f(\xi+h)-f(\xi)-hf'(\xi)}{h^2}.$$

It follows that

(53) $$\lim_{h\to 0}\lim_{k\to 0}\psi(\xi,h,k) = \tfrac{1}{2}[f''(\xi)],$$

if either of the limits exists

10. The generalized incremental ratio

Passing now to the nth derivative we similarly define

$$\phi(x,h_1,\dots,h_n) \equiv \frac{f(x+\sigma_n)-\sum f(x+\sigma_{n-1})+\sum f(x+\sigma_{n-2})-\dots(-)^n f(x)}{\prod\limits_1^n h_r},$$

where σ_r denotes the sum of any r of $h_1,\dots,h_n$ and the summation $\sum$ is taken for every distinct σ_r. We extend the theorems of § 9 to the nth derivative in the forms:

(54) *If, for every $r \leqslant n$, the repeated limit*

$$\lim_{h_1\to 0}\dots\lim_{h_r\to 0}\phi(\xi,h_1,\dots,h_r)$$

exists, then $f_n(\xi)$ exists and is equal to the limit

$$\lim_{h_1\to 0}\dots\lim_{h_n\to 0}\phi(\xi,h_1,\dots,h_n).$$

(55) *If $f_n(\xi)$ exists, then, for $r \leqslant n$, every repeated limit*

$$\lim_{h_1\to 0}\dots\lim_{h_r\to 0}\phi(\xi,h_1,\dots,h_r)$$

exists, and moreover $\phi(\xi,h_1,\dots,h_n)$ is convergent in $(h_1,\dots,h_n)$ at the origin, having the limit $f_n(\xi)$.

(56) *The continuity of $f_n(x)$ at $x = \xi$ implies and is implied by the continuity of $\phi(x,h_1,\dots,h_n)$ in $(x,h_1,\dots,h_n)$ at $(\xi,0,\dots,0)$.*

The proofs are the natural generalizations of those given for the second derivative. The second part of (55) may give difficulty. As an

indication of the procedure in the general case, consider $n = 3$. Write once more

$$F(t) = \frac{f(\xi+t)-f(\xi)}{t}.$$

Then

$$\phi(\xi, h_1, h_2, h_3) = \sum_{1,2,3} \frac{F(h_1+h_2+h_3)-F(h_1+h_2)-F(h_1+h_3)+F(h_1)}{h_2 h_3}$$
$$= \sum_{1,2,3} F''(h_1+\theta_2 h_2+\theta_3 h_3)$$

as in the proof of (49), where θ_2, θ_3 lie in $)0, 1($, provided that $F''(t)$ exists.

But by differentiation and rearrangement we have

$$F''(t) = 2t^{-3}\{f(\xi+t)-f(\xi)-f'(\xi)-\tfrac{1}{2}f''(\xi)\}-$$
$$-2t^{-2}\{f'(\xi+t)-f'(\xi)-tf''(\xi)\}+$$
$$+t^{-1}\{f''(\xi+t)-f''(\xi)\},$$

since $f''(x)$ exists.

As $t \to 0$, the several terms tend to the respective limits $\frac{1}{3}f'''(\xi)$, $-f'''(\xi), f'''(\xi)$. Hence, as h_1, h_2, h_3 tend simultaneously to 0 in any way, $\phi(\xi, h_1, h_2, h_3) \to f'''(\xi)$ and the convergence is established.

We generalized $\phi(x, h, k)$ for the nth derivative by forming $n-1$ successive incremental ratios of $\phi(x, h)$, all with regard to the argument x. If similarly we form $n-1$ successive incremental ratios all with regard to the argument h we generalize $\psi(x, h, k)$ in the form

$$\psi(x, h_1, \ldots, h_n) = Af(x) + \sum_{r=1}^{n} A_r f(x+h_r), \tag{57}$$

where $A, A_1, \ldots, A_n$ are the coefficients of partial fractions defined by the relation

$$\frac{1}{x\prod_1^n (x-h_r)} = \frac{A}{x} + \sum_{r=1}^{n} \frac{A_r}{x-h_r}.$$

It can be proved that

(58) $$n! \lim_{h_1 \to 0} \ldots \lim_{h_n \to 0} \psi(\xi, h_1, \ldots, h_n) = [f_n(\xi)],$$

if either limit exists.

More generally, we might consider a mixed incremental ratio of order n, in which the constituent incremental ratios are formed some with regard to the x-argument and some with regard to the h-argument. The order in which these are formed is immaterial, but the limit (if any) of the function as the increments are taken successively to zero in some order depends, in general, on this order. A variety of possible limits is thus obtainable but their discussion here is unprofitable.

If $f_n(\xi)$ exists, it will be found that all the nth-order incremental ratios are convergent as their increments tend to zero in any way, their limit being a numerical multiple of $f_n(\xi)$.

11. Schwarz's theorem

It is evident from the foregoing discussion that we cannot assert the existence of $f_n(\xi)$ from the convergence of $\phi(\xi, h_1,..., h_n)$ or of $\psi(\xi, h_1,..., h_n)$ along some special path to the origin. In particular, for $n = 2$, we are led to consideration of the respective limits

$$\lim_{h\to 0}\frac{f(\xi+h)-2f(\xi)+f(\xi-h)}{h^2} \tag{59}$$

and

$$\lim_{h\to 0}\frac{f(\xi+2h)-2f(\xi+h)+f(\xi)}{h^2}, \tag{60}$$

if we take $\phi(\xi, h, k)$ to the origin along the paths $h\pm k = 0$ or take $\psi(\xi, h, k)$ to the origin along the paths $h+k = 0$, $2h = k$. We cannot therefore expect to deduce the existence of $f''(\xi)$ from the existence of either of these limits alone, or indeed of the single limit of any function which does not involve $f'(x)$.

Examples readily confirm this belief. For take $\xi = 0$ and write

$$\left.\begin{aligned} &f(x) \equiv x\sin(x^{-2}) \qquad (x \neq 0) \\ &f(0) = 0. \end{aligned}\right\} \tag{61}$$

Then (59) exists, but neither (60) nor $f'(0)$ exists. Again write

$$f(x) = |x|. \tag{62}$$

Then (60) exists, but neither (59) nor $f'(0)$ exists. Lastly write

$$\left.\begin{aligned} &f(x) \equiv x\sin\{\pi\log_2(x^2)\} \qquad (x \neq 0) \\ &f(0) = 0. \end{aligned}\right\} \tag{63}$$

Then both (59) and (60) exist, but $f'(0)$ does not exist.

Yet it is not impossible for these limits, if they exist, to assume certain of the properties of $f''(x)$ and, of course, we can always use them for the direct calculation of $f''(x)$, if we otherwise know that it exists. There is, in this connexion, a theorem due to Schwarz, important in the theory of trigonometrical series:

(64) *If the limit* $$\lim_{h\to 0}\frac{f(x+h)-2f(x)+f(x-h)}{h^2}$$

exists and is zero everywhere in an interval throughout which $f(x)$ is continuous, then $f(x)$ is a linear function of x in the interval.

The conditions of continuity and evanescence of the limit are satisfied by any linear function $Ax+B$ and therefore also by any $f(x)-Ax-B$.

We can choose the disposable constants A, B, so that the new function vanishes at the end-points a, b of the interval. We have then

$$\phi(x) \equiv f(x) - \frac{x-b}{a-b}f(a) + \frac{x-a}{a-b}f(b).$$

Write further $\qquad \psi(x) \equiv \phi(x) + k^2(x-a)(x-b),$

where k remains a disposable constant. Then throughout (a, b)

$$\lim_{h\to 0} \frac{\psi(x+h) - 2\psi(x) + \psi(x-h)}{h^2} = 2k^2.$$

Hence, for every h less than some $\eta(x)$,

$$\psi(x+h) - 2\psi(x) + \psi(x-h) > 0.$$

At a maximum $x = \xi$ of $\psi(x)$

$$\psi(\xi \pm h) - \psi(\xi) < 0,$$

for every h less than some $\eta_1(\xi)$. Since these inequalities are contradictory, $\psi(x)$ cannot have a maximum in the interval. It is a continuous function and has a greatest value in the interval, which therefore must be at an end-point. At each of these the function vanishes. Hence throughout the interval (a, b) and for every k

$$\phi(x) + k^2(x-a)(x-b) \leqslant 0.$$

Take the limit of this inequality as $k \to 0$. Then throughout (a, b)

$$\phi(x) \leqslant 0.$$

But, by similar consideration of the function

$$\phi(x) - k^2(x-a)(x-b),$$

we show that throughout (a, b)

$$\phi(x) \geqslant 0.$$

Hence everywhere in the interval $\phi(x)$ vanishes and $f(x)$ is the linear function

$$\frac{x-b}{a-b}f(a) - \frac{x-a}{a-b}f(b).$$

Thus the theorem is proved.

WORKED EXAMPLE

If throughout some interval the n functions $y_1(x), \ldots, y_n(x)$ and their first N derivatives exist and if the Wronskian

$$(1, D)^{n-1}|y_1, \ldots, y_n|$$

vanishes throughout the interval, prove that every determinant

$$(D^a, D^b, \ldots)|y_1, \ldots, y_n| \qquad (0 \leqslant a, b, \ldots \leqslant N)$$

vanishes throughout the interval.†

† See also Chaundy, *Jour. London Math. Soc.* 8 (1933), 4–9, where the question is dealt with a little more fully.

If the vanishing of the Wronskian is sufficient to secure the linear dependence of $y_1,\ldots, y_n$, the vanishing of the other determinants naturally follows at once. Now this linear dependence fails to be secured by the vanishing of the Wronskian only in the neighbourhood of points ξ which are limiting points of zeros of the first minors of the Wronskian

$$(1, D,\ldots, D^{n-2})||y_1,\ldots, y_n||.$$

I shall therefore consider the rather wider theorem

(A) *If ξ is a limiting point of zeros of the Wronskian of $y_1,\ldots, y_n$ and if every $D^N y_n$ exists at ξ, then every determinant*

$$(D^a, D^b,\ldots)|y_1,\ldots, y_n| \qquad (0 \leqslant a, b,\ldots \leqslant N)$$

vanishes at ξ itself.

I shall use the symbol $(\xi)'$ to denote a set of points that has ξ as a limiting point but does not include ξ itself, so that in the above enunciation we might say simply that the Wronskian 'vanishes through $(\xi)'$'.

I shall prove theorem (A) inductively, showing that it is successively true for one, two, three,... functions y. It is an implied condition of the proof that we rely only on the *existence* of the final derivatives $D^N y$ and do not presume their continuity, still less the existence of any derivatives of higher order.

For a single function u, the theorem takes the form:

(1) *If $u = 0$ through some $(\xi)'$, then at ξ*

$$\textit{every } D^a u = 0 \qquad (0 \leqslant a \leqslant N).$$

Now, since u is differentiable, it is continuous, and therefore $u = 0$ at ξ, since $u = 0$ through $(\xi)'$. Hence

$$\frac{u(x)-u(\xi)}{x-\xi}$$

vanishes through $(\xi)'$ and therefore vanishes in the limit at $x = \xi$, i.e. $Du = 0$ at ξ.

Again, since u vanishes in every neighbourhood of ξ, then, by Rolle's theorem, Du also vanishes in every neighbourhood of ξ, i.e. Du vanishes through some $(\xi)'$. Hence, by induction, every $D^a u$ vanishes in some $(\xi)'$ and also at ξ up to the limits of differentiability. Thus (1) is established.

For a pair of functions u, v, theorem (A) takes the form:

(2) *If $(1, D)|u, v| = 0$ through some $(\xi)'$, then at ξ*

$$\textit{every } (D^a, D^b)|u, v| = 0 \qquad (0 \leqslant a, b \leqslant N).$$

I consider the possibilities:

(α) *One function, say v, vanishes through $(\xi)'$.*

Then, by (1), every $D^a v$ and therefore also every $(D^a, D^b)|u, v|$ vanishes at ξ.

(β) *One function, say v, has no zero in some closed interval $(\xi, \xi+h)$.*

Then throughout this interval the function $y = u/v$ exists and is differentiable up to $D^N y$. But

$$(1, D)|u, v| = v^2\, Dy,$$

and so, by the hypothesis of (2), Dy vanishes through $(\xi)'$, and therefore, by (1), every $D^a y$ $(1 \leqslant a \leqslant N)$ vanishes at ξ. If we expand $(D^a, D^b)|yv, v|$ by Leibniz's formula, the term in y disappears and we are left with a sum of multiples of $D^r y$ $(r \geqslant 1)$ all vanishing at ξ. Thus again every $(D^a, D^b)|u, v|$ vanishes at ξ.

(γ) *Both u, v vanish at ξ but nowhere else in $(\xi, \xi+h)$.*

Evidently every $(1, D^b)|u,v|$ vanishes at ξ, and there remain to be considered only $(D^a, D^b)|u,v|$ $(a, b \geqslant 1)$.

(γ_1) *Dv, say, vanishes through some $(\xi)'$.*

Then, as in (α), every $(D^a, D^b)|u,v|$ $(a, b \geqslant 1)$ vanishes at ξ.

(γ_2) *Dv, say, has no zero in some interval $)\xi, \xi+h)$, open at ξ.*

Then, throughout this interval, the function

$$Y \equiv \frac{(1, D)|u,v|}{Dv}$$

exists and is differentiable up to $D^{N-1}Y$. Actually we have, on reduction,

$$DY = \frac{v(D, D^2)|u,v|}{(Dv)^2}.$$

Now Y, and therefore also DY, vanishes through some $(\xi)'$; v does not vanish in some $)\xi, \xi+h)$ open at ξ. Thus $(D, D^2)|u,v|$ vanishes through some $(\xi)'$.

We cannot appeal to continuity to show that $(D, D^2)|u,v|$ vanishes also at ξ itself, since we might have $N = 2$, and we expressly do not presume the continuity of $D^N u, D^N v$. Consider then the limit, as $x \to \xi$, of

$$\frac{(1, D)|u,v|}{(x-\xi)^2}.$$

Near $x = \xi$ we have, by (35), since u, v vanish at ξ,

$$u = (x-\xi)(Du)_\xi + \tfrac{1}{2}(x-\xi)^2(D^2u)_\xi + \epsilon(x-\xi)^2,$$
$$Du = (Du)_\xi + (x-\xi)(D^2u)_\xi + \epsilon(x-\xi),$$

with similar approximations for v, Dv, and consequently

$$(1, D)|u,v| = \tfrac{1}{2}(x-\xi)^2\{(D, D^2)|u,v|\}_\xi + \epsilon(x-\xi)^2.$$

Thus
$$\lim_{x\to\xi} \frac{(1, D)|u,v|}{(x-\xi)^2} = \tfrac{1}{2}\{(D, D^2)|u,v|\}_\xi.$$

Since $(1, D)|u,v|$ vanishes throughout some $(\xi)'$, it follows that $(D, D^2)|u,v|$ vanishes at ξ.

We have now proved this much: if $(1, D)|u,v|$ vanishes through some $(\xi)'$, then *either* every $(D^a, D^b)|u,v|$ $(0 \leqslant a, b \leqslant N)$ vanishes at ξ, *or else* every $(1, D^b)|u,v|$ $(1 \leqslant b \leqslant N)$ vanishes at ξ, and, in addition, $(D, D^2)|u,v|$ vanishes through some $(\xi)'$ and also at ξ. Thus, if $N = 2$, the theorem is complete. If $N = 3$, there are still certain $(D^a, D^b)|u,v|$ $(a, b \geqslant 1)$ outstanding. Write these as

$$(D^a, D^b)|Du, Dv| \qquad (a, b \geqslant 0).$$

But
$$(1, D)|Du, Dv| = (D, D^2)|u,v|$$

vanishes in some $(\xi)'$. Hence the completion of the proof for $N = 3$ follows from that for $N = 2$. We can proceed inductively in this way to any N, and the theorem is established.

From (2) we proceed inductively to the theorem for three variables:

(3) *If $(1, D, D^2)|u,v,w| = 0$ through some $(\xi)'$, then at ξ*
every $(D^a, D^b, D^c)|u,v,w| = 0$ $\qquad (0 \leqslant a, b, c \leqslant N)$.

The proof proceeds along similar lines to that of (2). It will be sufficient to indicate the modifications due to the presence of the third function.

(α) If one function, say w, vanishes throughout some $(\xi)'$, the theorem is obvious as before.

(β) If one function, say w, has no zero in some closed interval $(\xi, \xi+h)$, we write $u \equiv yw$, $v \equiv zw$. Then $(1, D)|Dy, Dz|$ vanishes through some $(\xi)'$ and we can write $(D^a, D^b, D^c)|u, v, w|$ as a sum of multiples of $(D^p, D^q)|Dy, Dz|$. Thus (3) follows from (2).

(γ) If u, v, w all vanish at ξ but nowhere else in some $(\xi, \xi+h)$, we have at once that every $(1, D^a, D^b)|u, v, w| = 0$ at ξ.

(γ_1) If some $(D, D^2)|v, w|$ vanishes through some $(\xi)'$, then, by (2), every $(D^a, D^b)|v, w|$ $(a, b \geqslant 1)$ vanishes at ξ, and therefore also every

$$(D^a, D^b, D^c)|u, v, w| \qquad (a, b, c \geqslant 1).$$

(γ_2) If $(D, D^2)|v, w|$ has no zero in a half-open interval $)\xi, \xi+h)$, we consider the function

$$Y = \frac{(1, D, D^2)|u, v, w|}{(D, D^2)|v, w|},$$

which has the derivative

$$DY = \frac{(1, D)|v, w|\,(D, D^2, D^3)|u, v, w|}{\{(D, D^2)|v, w|\}^2},$$

vanishing through some $(\xi)'$. If $(1, D)|v, w|$ vanishes through $(\xi)'$, then, by (2), every $(D^a, D^b)|v, w|$ $(a, b > 0)$ vanishes at ξ, and therefore also every

$$(D^a, D^b, D^c)|u, v, w| \quad (a, b, c > 0).$$

Alternatively $(D, D^2, D^3)|u, v, w|$ vanishes through some $(\xi)'$. To show that $(D, D^2, D^3)|u, v, w|$ vanishes actually at ξ we appeal to the limit

$$\lim_{x \to \xi} \frac{(1, D, D^2)|u, v, w|}{(x-\xi)^3} = \frac{1}{3!}\{(D, D^2, D^3)|u, v, w|\}_{x=\xi}.$$

The proof of (3) is then completed like the proof of (2).

The extension of the argument to the case of n variables introduces no new difficulties and we may therefore regard theorem (A) as proved and, as a consequence, the original theorem which it includes.

EXAMPLES V

1. Prove the formulae

$$D^n \frac{1}{x^2+1} = (-)^n n! \sin^{n+1}\theta \sin(n+1)\theta,$$

$$D^n \frac{x}{x^2+1} = (-)^n n! \sin^{n+1}\theta \cos(n+1)\theta,$$

where $\theta = \cot^{-1}x$.

If $y = (x^2+2x\cos\alpha+1)^{-1}$, prove that

$$\frac{1}{y^{n+1}}\frac{d^n y}{dx^n} = (-)^n n! \left\{(n+1)x^n + \frac{(n+1)n}{2!}x^{n-1}\frac{\sin 2\alpha}{\sin\alpha} + \dots + \frac{\sin(n+1)\alpha}{\sin\alpha}\right\}.$$

2. Obtain formulae for the nth derivatives of

$$1/(x^2-ax+a^2), \qquad x/(x^2-ax+a^2), \qquad 1/(x^3+a^3).$$

3. Obtain the nth derivative of $(px+q)/(ax^2+2bx+c)$ in the form

$$(-)^n n! \frac{a^{\frac{1}{2}n}(aq^2-2bpq+cp^2)^{\frac{1}{2}}}{(ac-b^2)^{\frac{1}{2}}(ax^2+2bx+c)^{\frac{1}{2}(n+1)}} \sin\left\{(n+1)\cot^{-1}\frac{ax+b}{\sqrt{(ac-b^2)}} + \cot^{-1}\frac{aq-bp}{p\sqrt{(ac-b^2)}}\right\},$$

where $ac-b^2 > 0$.

4. If $U_n = D^n \dfrac{px+q}{ax^2+2bx+c}$, prove that

$$(ax^2+2bx+c)U_{n+2}+2(n+2)(ax+b)U_{n+1}+(n+1)(n+2)aU_n = 0.$$

5. Prove that, if k be any real number and $D^n u$ exist, then

$$D^n(u^k) = n!\sum \frac{k(k-1)\dots(k-\sigma+1)}{\alpha!\beta!\dots\theta!}u^{k-\sigma}(Du)^\alpha\left(\frac{D^2u}{2!}\right)^\beta\dots\left(\frac{D^nu}{n!}\right)^\theta,$$

where the summation $\sum$ is taken over all positive integral or zero values of $\alpha, \beta, \dots, \theta$ which satisfy the equation $\alpha+2\beta+\dots+n\theta = n$, and σ is written for $\alpha+\beta+\dots+\theta$.

6. Prove that, granted differentiability,

$$D^n\phi(u) = A_1\phi'(u)+\frac{A_2}{2!}\phi''(u)+\dots+\frac{A_n}{n!}\phi_n(u),$$

where A_r is independent of the form of ϕ.

Prove that

(i) $A_r = D^n(u^r)-(r, 1)uD^n(u^{r-1})+(r, 2)u^2D^n(u^{r-2})-\dots+(-)^{r-1}ru^{r-1}D^n(u)$,

where (r, s) denotes a binomial coefficient [Bertrand.]

(ii) $A_r = \sum \dfrac{n!r!}{\alpha!\beta!\dots\theta!}(Du)^\alpha\left(\dfrac{D^2u}{2!}\right)^\beta\dots\left(\dfrac{D^nu}{n!}\right)^\theta$,

where the summation $\sum$ is taken over all positive integral or zero values of $\alpha, \beta, \dots, \theta$ which satisfy the equations

$$\alpha+\beta+\dots+\theta = r, \qquad \alpha+2\beta+\dots+n\theta = n.$$

7. If
$$D^n\phi(e^x) = A_1\phi'(e^x)+\frac{A_2}{2!}\phi''(e^x)+\dots+\frac{A_n}{n!}\phi_n(e^x),$$

show that

$$A_r = e^{rx}\{r^n-(r, 1)(r-1)^n+(r, 2)(r-2)^n-\dots+(-)^{r-1}r\};$$

and if
$$(-D)^n\frac{1}{e^x-1} = \frac{B_1}{e^x-1}+\frac{B_2}{(e^x-1)^2}+\dots+\frac{B_{n+1}}{(e^x-1)^{n+1}},$$

show that

$$B_r = r^n-(r-1, 1)(r-1)^n+(r-1, 2)(r-2)^n-\dots+(-1)^{r-1},$$

where (r, s), etc., denote binomial coefficients.

8. Prove that, subject to considerations of differentiability,

(i) $D^n\phi(x^{-1}) = (-)^n\displaystyle\sum_{r=0}^{n-1}\frac{n(n-1)^2(n-2)^2\dots(n-r-1)^2(n-r)}{r!}x^{r-2n}\phi_{n-r}(x^{-1})$,

(ii) $D^n\phi(x^2) = \displaystyle\sum_{r=0}\frac{n!}{(n-2r)!\,r!}(2x)^{n-2r}\phi_{n-r}(x^2)$,

the last term in $\sum$ being given by $r = \frac12(n-1)$ or $\frac12 n$ according as n is odd or even.

9. Prove that

(i) $D^{2n}(1\pm x^2)^{n-\frac12} = (\pm)^n 1^2 . 3^2 \dots (2n-1)^2(1\pm x^2)^{-(n+\frac12)}$,

(ii) $(xD^{2n+1}-D^{2n})(1-x^2)^{n+\frac12} = (-)^{n+1}1^2.3^2\dots(2n-1)^2(2n+1)(1-x^2)^{-(n+\frac12)}$.

10. Prove that, if $x = \cos u$,

$$\text{(i)}\quad D^n(1-x^2)^{n+\frac{1}{2}} = (-)^n \frac{1.3\ldots(2n+1)}{n+1}\sin(n+1)u,$$

$$\text{(ii)}\quad D^n(1-x^2)^{n-\frac{1}{2}} = (-)^n 1.3\ldots(2n-1)\frac{\cos nu}{\sin u},$$

$$\text{(iii)}\quad D^n\{x(1-x^2)^{n-\frac{1}{2}}\} = (-)^n 1.3\ldots(2n-1)\frac{\cos(n+1)u}{\sin u}.$$

11. Prove that, if $x = \sinh u$,

$$D^n(1+x^2)^{n+\frac{1}{2}} = \frac{1.3\ldots(2n+1)}{n+1}U,$$

where U is $\sinh(n+1)u$ or $\cosh(n+1)u$ according as n is odd or even.

12. If m, n are positive integers such that $m > n$, prove, by Taylor's theorem or otherwise, that

$$\frac{D^{m-n}}{(m-n)!}(x^2+a)^m = (x^2+a)^n \frac{D^{m+n}}{(m+n)!}(x^2+a)^m.$$

13. If x, y, u, v are functions of t which satisfy the relations

$$xy = uv, \qquad x_1 y_1 = u_1 v_1,$$

prove that, for every n, $\qquad x_n y_n = u_n v_n,$

the suffixes denoting differentiation in t and the existence of the derivatives being presumed.

14. If x, y, u, v are functions of t which satisfy the relations

$$x_1 y - x y_1 = u_1 v - u v_1, \qquad x_2 y_1 - x_1 y_2 = u_2 v_1 - u_1 v_2,$$

prove that, for every n,

$$x_n y_{n-1} - x_{n-1} y_n = u_n v_{n-1} - u_{n-1} v_n,$$

and more generally, for every r, s,

$$x_r y_s - x_s y_r = u_r v_s - u_s v_r,$$

where the suffixes denote differentiation in t and the existence of the derivatives is presumed.

15. Show that $\qquad y = C_1 e^{C_2 x} + C_3 e^{C_4 x} + \ldots + C_{2n-1} e^{C_{2n} x}$

satisfies the differential equation

$$(D, 1)^n |(D, 1)^n y| = 0.$$

16. If y, $y_1, \ldots, y_n$ be any $n+1$ differentiable functions of x, prove that

$$\text{(i)}\quad (1, D)^{n-1}|yy_1, \ldots, yy_n| = y^n (1, D)^{n-1}|y_1, \ldots, y_n|,$$

$$\text{(ii)}\quad (1, D_x)^{n-1}|y_1, \ldots, y_n| = \left(\frac{dy}{dx}\right)^{\frac{1}{2}n(n-1)} (1, D_y)^{n-1}|y_1, \ldots, y_n|.$$

17. (i) Prove that $\qquad (1, D)^n |(1, y)^n| = n!!(Dy)^{\frac{1}{2}n(n+1)},$

where $n!!$ denotes $n!(n-1)!\ldots 2!1!$.

(ii) If

$$u_n = \frac{(1, D)^n |(1, y)^{n-1}, x|}{(1, D)^n |(1, y)^n|},$$

show that

$$(n+1)u_{n+1} = \frac{du_n}{dy},$$

and prove that, if $n > 1$,

$$\frac{d^n x}{dy^n} = \frac{(-1)^{n-1}(D^2, D^3,\dots, D^n)|y, y^2,\dots, y^{n-1}|}{(n-1)!!(Dy)^{\frac{1}{2}n(n+1)}}.$$

18. If $Y_1,\dots, Y_n$ denote the minors of $D^{n-1}y_1,\dots, D^{n-1}y_n$ in

$$\Delta = (1, D)^{n-1}|y_1,\dots, y_n|,$$

prove that $$(1, D)^{n-1}|Y_1,\dots, Y_n| = \Delta^{n-1},$$

and, more generally, that

$$(1, D)^{r-1}|Y_1,\dots, Y_r| = (-)^{r(n-1)}\Delta^{r-1}(1, D)^{n-r-1}|y_{r+1},\dots, y_n|.$$

19. If, throughout some interval, $D^n z$ exists and the Wronskian of $y_1,\dots, y_n$ vanishes, prove that the Wronskian of $z, y_1,\dots, y_n$ vanishes throughout the interval.

20. In the exceptional case where $y_1,\dots, y_n$ are not linearly connected throughout an interval (a, b), although the Wronskian $(1, D)^{n-1}|y_1,\dots, y_n|$ vanishes throughout (a, b), show that the minors

$$(1, D)^{n-2}||y_1,\dots, y_n||$$

must all have a common zero in (a, b).

21. Prove that

$$D^n\{(1, D)|u, v|\} = \sum_{r=0} \frac{n!(n-2r-1)}{(n-r+1)!r!}(D^r, D^{n-r+1})|u, v|,$$

where the last term on the right is given by $r = \frac{1}{2}(n-1)$ or $\frac{1}{2}n$, according as n is odd or even. Prove more generally that

$$D^n\{(1, D)^{p-1}|u_1,\dots, u_p|\}$$
$$= \sum \frac{n!\prod (a_r - a_s - r - s)}{a_1!(a_2+1)!\dots(a_r+r-1)!\dots(a_p+p-1)!}(D^{a_1},\dots, D^{a_p})|u_1,\dots, u_p|,$$

where $\sum$ is taken over all positive integers $a_1,\dots, a_p$ such that $a_1 < a_2 < \dots < a_p$ and $a_1 + a_2 + \dots + a_p = n$, and $\prod$ is taken over every pair of suffixes r, s such that $r > s$.

22. If $$Y_s = (1, D)^{n-r}|y_s, y_{r+1},\dots, y_n| \qquad (s = 1,\dots, r),$$

prove that

$$(1, D)^{r-1}|Y_1,\dots, Y_r| = (1, D)^{n-1}|y_1,\dots, y_n|.\{(1, D)^{n-r-1}|y_{r+1},\dots, y_n|\}^{r-1}.$$

23. (i) If $\theta = (1, D)^n|v, u_1,\dots, u_n|, \quad \phi = (1, D)^{n-1}|Du_1,\dots, Du_n|,$

prove that

$$\theta'\phi - \theta\phi' = (1, D)^n|Dv, Du_1,\dots, Du_n|.(1, D)^{n-1}|u_1,\dots, u_n|,$$

accents denoting differentiation.

(ii) If $\theta = (1, D)^{n+1}|v_1, v_2, u_1,\dots, u_n|, \quad \phi = (1, D)^{n-1}|D^2u_1,\dots, D^2u_n|,$

prove that

$$\theta'\phi - \theta\phi' = (1, D^2, D^3,\dots, D^{n+2})|v_1, v_2, u_1,\dots, u_n|.(1, D)^{n-1}|Du_1,\dots, Du_n| -$$
$$-(1, D^2, D^3,\dots, D^n)|u_1,\dots, u_n|.(1, D)^{n+1}|Dv_1, Dv_2, Du_1,\dots, Du_n|.$$

(iii) Obtain the corresponding expression for $\theta'\phi - \theta\phi'$, where

$$\theta \equiv (1, D)^{n+r-1}|v_1,\dots, v_r, u_1,\dots, u_n|, \qquad \phi \equiv (1, D)^{n-1}|D^r u_1,\dots, D^r u_n|.$$

24. If, for some fixed n and for every $r \leqslant n$,

$$\theta_r = (1, D)^{n-r}|D^{2r}u, D^{2r+2}u,\dots, D^{2n}u|,$$

where u is a polynomial of degree not exceeding $2n+1$, prove that

$$\theta'_{r+1}\theta_{r-1}-\theta_{r+1}\theta'_{r-1} = \theta_r^2,$$

where accents denote differentiation.

Show that $\theta_{r+1}^2/\theta_r^2$ and $\theta_r^2/\theta_{r+1}^2$ are each derivatives of rational functions and that

$$\theta''_{r+1}\theta_r-2\theta'_{r+1}\theta'_r+\theta_{r+1}\theta''_r = 0.$$

25. If $u(x)$, $v(x)$ are functions of x differentiable sufficiently often, show that

$$(1,D)^n|(1,x)^n u| = n!!u^{n+1},$$

$$(1,D)^n|(1,x)^{n-1}u, v| = (n-1)!!u^{n+1}D^n(v/u),$$

$$(D,D^2,\dots,D^n)|(1,x)^{n-1}u| = (-)^n(n-1)!!u^{n+1}D^n(1/u).$$

26. If $\phi_1(x),\dots,\phi_m(x)$ are m linearly independent polynomials of order less than m and $\psi_1(x),\dots,\psi_n(x)$ are n linearly independent polynomials of order less than n, prove that the eliminant of the polynomials $f(x)$, $g(x)$ of respective orders m, n may be written, save for a numerical factor,

$$\frac{(D,1)^{m+n-1}|\psi_1 f,\dots,\psi_n f,\phi_1 g,\dots,\phi_m g|}{(D,1)^{m-1}|\phi_1,\dots,\phi_m|.(D,1)^{n-1}|\psi_1,\dots,\psi_n|}.$$

27. With the notation of § 4 prove that the leading coefficients in $f_s(x)$, $g_s(x)$ are respectively $\quad (-)^{m-s-1}a_0 C_s, \qquad (-)^{m-s-1}b_0 C_s,$

where a_0, b_0, C_s are the leading coefficients respectively in $f(x)$, $g(x)$, $R_s(x)$.

28. With the notation of § 4 and of the preceding example prove that

$$f_s(x)g_{s+1}(x)-f_{s+1}(x)g_s(x) = (-)^{n-s}C_s^2.$$

29. If $f(x)$ vanishes at $c_1,\dots,c_n$ in (a,b) and if $f_n(x)$ exists throughout the interval, prove that, at any point x of the interval,

$$f(x) = (x-c_1)\dots(x-c_n)\frac{f_n(\xi)}{n!},$$

where $\xi = \xi(x)$ is some other point of the interval.

30. If $\theta(n,a,h)$ is defined by Taylor's formula

$$f(a+h) = \sum_{r=0}^{n-1}\frac{h^r}{r!}f_r(a)+\frac{h^n}{n!}f_n(a+\theta h),$$

prove that, if $f(x)$ is (i) e^x, (ii) $\log x$, (iii) x^{n+k} (k a positive integer), then $\theta(n,a,h)$ is

$$\text{(i)}\ \ h^{-1}\log\left\{1+\frac{h}{n+1}+\frac{h^2}{(n+1)(n+2)}+\frac{h^3}{(n+1)(n+2)(n+3)}+\dots\text{ to }\infty\right\};$$

$$\text{(ii)}\ \ -\frac{a}{h}+\frac{a}{h}\left\{1-\frac{n}{n+1}\left(\frac{h}{a}\right)+\frac{n}{n+2}\left(\frac{h}{a}\right)^2-\dots\text{ to }\infty\right\}^{-1/n};$$

$$\text{(iii)}\ \ -\frac{a}{h}+\frac{a}{h}\left\{1+\frac{k}{n+1}\left(\frac{h}{a}\right)+\frac{k(k-1)}{(n+1)(n+2)}\left(\frac{h}{a}\right)^2+\dots+\frac{k!n!}{(n+k)!}\left(\frac{h}{a}\right)^k\right\}^{1/k}.$$

31. Show that, if x is any real number,

$$\{2x-2\log(1+x)\}^{-\frac{1}{2}}-x^{-1}$$

and

$$\{3\log(1+x)-3x+\tfrac{3}{2}x^2\}^{-\frac{1}{3}}-x^{-1}$$

both lie between 0, 1.

32. In the formula of the mean (28), prove that,

(i) if $f_{n+1}(a)$ exists and is not zero, then, as $h \to 0$,

$$\theta \to (n+1)^{-1};$$

(ii) if $f_{n+1}(a),\dots,f_{n+r-1}(a)$ are all zero, but $f_{n+r}(a)$ exists and is not zero, then, as $h \to 0$,

$$\theta \to (n+r\,!\,r)^{-1/r}.$$

33. In the fractional formula of the mean (27), prove that, if $f_{n+1}(a)$, $g_{n+1}(a)$ exist, and, if

$$f_{n+1}(a)\,g_n(a) \neq g_{n+1}(a)\,f_n(a),$$

then, as $h \to 0$,

$$\theta \to (n+1)^{-1}.$$

34. If, in the linear formula of the mean (28), θ is independent of h throughout some interval of values of a, h, show that it is also independent of a and that $f(x)$ is a polynomial in x of degree not exceeding $n+1$.

[You may assume the existence of as many derivatives $f_{n+r}(x)$ as you please.]

35. If, in the linear formula of the mean, $f(x) = \exp(-x^{-2})$, $a = 0$, show that $\theta \to 1$ as $h \to 0$, whatever n.

36. Show that the linear formula of the mean (28) applies to the case

$$f(x) = x^n(\log x)^p \qquad (a = 0),$$

where n is the positive integer n of the formula and p is any positive number, although $f_n(0)$ does not exist. Show, however, in this case that

$$\lim_{h\to 0}\theta = e^{-(1+\frac{1}{2}+\dots+1/n)}.$$

37. Discuss the applicability of the fractional formula of the mean (27) to the functions

$$f(x) = 4x^6+4x^5-5x^4, \qquad g(x) = 8x^5+10x^4-20x^3$$

for the case $a = 0$, $h = 1$, $n = 3$, showing that

$$f_3(x) = xg_3(x)$$

but

$$f(1)-f(0) = -\tfrac{3}{2}[g(1)-g(0)].$$

38. Show that under certain conditions

$$\frac{f(a+h)-\sum\limits_{r=0}^{m-1}\dfrac{h^r}{r!}f_r(a)}{g(a+h)-\sum\limits_{r=0}^{n-1}\dfrac{h^r}{r!}g_r(a)} = \frac{(n-1)!\,[h(1-\theta)]^{m-n}f_m(a+\theta h)}{(m-1)!\,g_n(a+\theta h)}.$$

39. Prove that under certain conditions

$$f(a+h)-\sum_{r=0}^{n-1}\frac{h^r}{r!}f_r(a) = \frac{h^n f_{n-1}(a)f_n(a+\theta h)}{n!\,f_{n-1}(a+\theta h)-(n-1)!\,(1-\theta)hf_n(a+\theta h)}$$

for some θ in $)0,1($.

40. Prove the generalized formula for the remainder in Taylor's series

$$f(a+h)-\sum_{r=0}^{n-1}\frac{h^r}{r!}f_r(a) = \frac{(1-\theta)^{n-1}h^{n-1}\{\psi(h)-\psi(0)\}f_n(a+\theta h)}{(n-1)!\,\psi'(\theta h)},$$

where $\psi(t)$ is any function such that $\psi'(t)$ exists and does not vanish in $)0,h($.

Deduce the particular forms of remainder:

(i) $\dfrac{(1-\theta)^{n-p-1}h^n}{(p+1)(n-1)!}f_n(a+\theta h)$ [SCHLOMLICH-ROCHE],

(ii) $\left(\dfrac{1}{\theta}-1\right)^{n-1}\dfrac{h^n}{n!}f_n(a+\theta h)$,

(iii) $\dfrac{(1-\theta)^{n-1}h^n}{(p+1)(n-1)!\theta^p}f_n(a+\theta h)$,

(iv) $\dfrac{(1-\theta)^{n-1}h^{n-1}}{(n-1)!}2e^{(\frac{1}{2}-\theta)h}\sinh\tfrac{1}{2}h\, f_n(a+\theta h)$

41. If Δ is the algebraic operator that changes a into $a+1$, prove that

$$\frac{\Delta^n f(a)}{\Delta^n g(a)}=\frac{f_n(a+n\theta)}{g_n(a+n\theta)} \qquad (0<\theta<1)$$

If $f'(x), f''(x), \dots, f_n(x)$ are all positive in $(a, a+n)$, prove that $\Delta f(a), \Delta^2 f(a), \dots, \Delta^n f(a)$ are all positive.

42. If $f(x)$ and $x^n+a_1x^{n-1}+\dots+a_n$ are two polynomials which have all their zeros real and distinct, show that the same is true of the polynomial

$$f(x)+a_1f'(x)+\dots+a_nf_n(x).$$

43. If $f(x)$, $g(x)$ be polynomials of which $g(x)$ has a repeated factor $(x-a)^n$, show that, when $f(x)/g(x)$ is put into partial fractions, the repeated factor $(x-a)^n$ gives rise to the set of n partial fractions

$$\sum_{r=0}^{n}\frac{A_r}{(x-a)^{n-r}}, \quad \text{where} \quad A_r=\frac{1}{r!}\left[\frac{d^r}{dx^r}\left\{\frac{(x-a)^n f(x)}{g(x)}\right\}\right]_{x=a}.$$

Show that A_r may also be written in either of the forms

(i) $\dfrac{(n-1)!(r-1)!}{(n+r)!\{g_n(a)/n!\}^{r+1}}(D^n, D^{n+1}, \dots, D^{n+r})|(1, x-a)^{r-1}g(x), (x-a)^n f(x)|_{x=a}$,

(ii) $(r-1)!\begin{vmatrix} g_n(a)/n! & 0 & \cdot & f(a) \\ g_{n+1}(a)/(n+1)! & g_n(a)/n! & \cdot & f_1(a)/1! \\ \cdot & \cdot & \cdot & \\ g_{n+r}(a)/(n+r)! & g_{n+r-1}(a)/(n+r-1)! & \cdot & f_r(a)/r! \end{vmatrix} \Big/ \left\{\dfrac{g_n(a)}{n!}\right\}^{r+1}$.

44. If u, v be sufficiently differentiable functions of x, prove that

$$D^m(uv^{-n})$$
$$=(-)^m v^{-(m+n)}\begin{vmatrix} u & Du & \dfrac{D^2u}{2!} & \cdot & \dfrac{D^mu}{m!} \\ mv & (m+n-1)Dv & (m+2n-2)\dfrac{D^2v}{2!} & \cdot & nm\dfrac{D^mv}{m!} \\ 0 & (m-1)v & (m+n-2)Dv & \cdot & n(m-1)\dfrac{D^{m-1}v}{(m-1)!} \\ 0 & 0 & (m-2)v & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{vmatrix}$$

If u, v be polynomials having neither common nor repeated factors, then uv^{-n} is the exact derivative of a rational algebraic function $\phi(x)v^{-(n-1)}$, if and only if v is a factor of the polynomial

$$\begin{vmatrix} u & Du & \dfrac{D^2u}{2!} & \cdot & \dfrac{D^{n-1}u}{(n-1)!} \\ (n-1)Dv & (2n-2)\dfrac{D^2v}{2!} & (3n-3)\dfrac{D^3v}{3!} & & n(n-1)\dfrac{D^nv}{n!} \\ 0 & (n-2)Dv & (2n-3)\dfrac{D^2v}{2!} & & n(n-2)\dfrac{D^{n-1}v}{(n-1)!} \\ 0 & 0 & (n-3)Dv & & \\ \cdot & \cdot\ \cdot\ \cdot & & \cdot\ \cdot & \end{vmatrix}$$

45 In the notation of § 9 (52) prove that $f''(\xi)$ exists, if $\psi(\xi, h, k)$ is continuous in (h, k) at $(0, 0)$ and $f'(x)$ exists near $x = \xi$

46. If n, r are positive integers, and if

$$\left.\begin{aligned} f(x) &= x^{n(r+1)}\sin(x^{-r}) \qquad (x \neq 0) \\ f(0) &= 0 \end{aligned}\right\},$$

prove that only the first n derivatives of $f(x)$ exist at $x = 0$, but that the umbral derivatives $[f_{n+1}(0)], \ldots, [f_{nr+n-1}(0)]$ also exist

47 If $\phi(x) = f_m(x)$, $\psi(x) = f_{m-p}(x)$, where $f_m(x)$ exists near $x = \xi$, and if the umbral derivative $[\phi_n(\xi)]$ exists, prove that the umbral derivative $[\psi_{n+p}(\xi)]$ exists and that

$$[\psi_{n+p}(\xi)] = [\phi_n(\xi)]$$

48 If n is a positive integer and if

$$\left.\begin{aligned} f(x) &= x^{n+1/(n-2)}\sin\{x^{-1/(n-2)}\} \qquad (x \neq 0) \\ f(0) &= 0 \end{aligned}\right\},$$

prove that $[f_n(0)]$ exists, but that if $\phi(x) = f_r(x)$, then $[\phi_{n-r}(0)]$ does not exist for $r = 1, 2, \ldots$.

49. In generalization of Schwarz's theorem prove that,

(i) if, at every point of an interval, $f(x)$ is continuous and

$$\lim_{h\to 0}\frac{\sum p_r\{f(x+a_rh)-f(x)\}}{h^2} = 0,$$

where every p_r is positive and $\sum p_r a_r = 0$, then $f(x)$ is a linear function of x in the interval,

(ii) if the umbral second derivative $[f''(x)]$ vanishes throughout an interval, then $f(x)$ is a linear function of x in the interval

50. (i) If, at every point an interval,

$$\lim_{h\to 0}\frac{f(x+2h) - 2f(x+h) + f(x)}{h^2} = 0,$$

show that the continuity of $f(x)$ is insufficient but the differentiability of $f(x)$ is sufficient to secure that $f(x)$ be linear in the interval.

(ii) More generally show that, if, at every point of an interval,

$$\lim_{h\to 0}\frac{qf(x+ph) - pf(x+qh) + (p-q)f(x)}{h^2} = 0,$$

where p, q have the same sign, then the continuity of $f(x)$ is insufficient but the differentiability is sufficient to secure that $f(x)$ be linear in the interval.

51. If, at every point of an interval,

$$\lim_{h\to 0}\frac{f(x+3h)-3f(x+h)+3f(x-h)-f(x-3h)}{h^3}=0,$$

show that continuity is insufficient to secure that $f(x)$ be a quadratic function throughout the interval.

If $f'(x)$ exists throughout the interval, show that the above condition is equivalent to the condition

$$\lim_{h\to 0}\frac{f(x+h)-f(x-h)-2hf'(x)}{h^3}=0$$

throughout the interval.

If, in addition, $f'(x)$ is continuous throughout the interval, show that the above condition secures that $f(x)$ be a quadratic function throughout the interval.

52. If, at every point of an interval,

$$\lim_{h\to 0}\frac{f(x+2h)-4f(x+h)+6f(x)-4f(x-h)+f(x-2h)}{h^4}=0,$$

show that the existence and continuity of $f'(x)$ is not enough to secure that $f(x)$ be a cubic function of x throughout the interval.

If $f''(x)$ exists throughout the interval, show that the above condition is equivalent to the condition

$$\lim_{h\to 0}\frac{f(x+h)+f(x-h)-2f(x)-h^2f''(x)}{h^4}=0$$

throughout the interval.

If, in addition, $f''(x)$ is continuous throughout the interval, show that the above condition secures that $f(x)$ be a cubic function throughout the interval.

VI

PARTIAL DIFFERENTIATION

1. The nature of partial differentiation

We have so far restricted the notion of differentiability and the processes of differentiation to functions of a single variable. We proceed to extend them to functions of many variables.

Consider first a function of two variables, say $f(x,y)$. By fixing a value η of y we reduce it to $f(x,\eta)$, a function of one variable x only. We can then discuss its differentiability at $x = \xi$ by studying the limit

$$\lim_{h\to 0}\frac{f(\xi+h,\eta)-f(\xi,\eta)}{h}. \tag{1}$$

On the other hand, we can similarly fix x at the value ξ and discuss the differentiability at $y = \eta$ of $f(\xi,y)$, a function of y only, by studying the limit

$$\lim_{k\to 0}\frac{f(\xi,\eta+k)-f(\xi,\eta)}{k}. \tag{2}$$

We can satisfy ourselves that these limits are independent by considering such a function as

$$f(x,y) = \phi(x)+\psi(y).$$

We may, therefore, speak of the differentiability of $f(x,y)$ in x or in y independently, just as we speak of its continuity in x or in y independently.

The limits (1) and (2), when they exist, are known as the *partial derivatives* of $f(x,y)$ in x and in y respectively,† and the process of obtaining them is known as *partial differentiation*. For partial derivatives we introduce a distinguishing form of the notation for ordinary derivatives and write

$$\frac{\partial f}{\partial x} = \lim_{h\to 0}\frac{f(x+h,y)-f(x,y)}{h}, \text{ etc.}$$

In analogy with the notation $f'(x)$ for the derivative of $f(x)$ we may write

$$f_x(x,y), \qquad f_y(x,y)$$

for the partial derivatives of $f(x,y)$, and

$$f_1(x_1,x_2), \qquad f_2(x_1,x_2),$$

† 'With respect to x' is, of course, the orthodox wording.

for the partial derivatives of $f(x_1, x_2)$, or more generally for the partial derivatives of f in its first and its second argument, respectively.

The operation of partial differentiation is represented by the partial differential operators

$$\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \quad \text{or} \quad \partial_x, \ \partial_y.$$

The essential principle of partial differentiation is that, in turn, all but one of the arguments of the function are ignored, i.e. are treated as constants. There is, of course, no novelty in this, for we have already been accustomed to differentiate specific functions, e.g. $(x+a)^n$, which involve constants. The interest of the present chapter lies in the exploration of the theoretical consequences of this principle.

2. Differentiability of $f(x, y)$

If we take the geometric point of view, as we did in discussing the continuity of $f(x, y)$, and regard $f(x, y)$ as a function of position in the plane of (x, y), we find that the mere existence, at some point, of $f_x(x, y), f_y(x, y)$ does not carry us very far. By chapter IV (5), it secures the continuity of $f(x, y)$ in x and in y respectively. But that, by chapter III § 10, is insufficient to secure continuity in (x, y) at the point. We can go farther and show by an example that

(3) *The existence of $f_x(x, y)$, $f_y(x, y)$ at a point does not secure the continuity of $f(x, y)$ there.*

For write $$f(x, y) \equiv \frac{xy}{x^2+y^2}, \qquad f(0,0) = 0. \tag{4}$$

Then $$f(x, 0) = 0 = f(0, y)$$

and so $f_x(0,0)$, $f_y(0,0)$ both exist (being zero). But this function (4) is the function (23) of chapter III, which is discontinuous at $(0,0)$. It converges in fact to any limit $k/(1+k^2)$ in the interval $(-\frac{1}{2}, \frac{1}{2})$, if we travel to the origin along the appropriate path $y = kx$.

An adequate condition of differentiability at a point ought not merely to secure continuity at the point. It ought, from our geometric point of view, to secure differentiability along every curve through the point, and it ought to survive a change of coordinates. In analytical terms, our generalized definition of differentiability must be consistent with the simple definition for a substitution $x = x(t)$, $y = y(t)$, and must be self-consistent for the wider substitution $x = x(u, v), y = y(u, v)$, the substituted functions $x(t)$, $y(t)$, $x(u, v)$, $y(u, v)$ being, of course, them-

selves differentiable. Our requirements, in precise language, are respectively:

(5) *If $f(x,y)$ is differentiable at $\{x(t_0), y(t_0)\}$ and $x(t)$, $y(t)$ are differentiable at t_0, then $f\{x(t), y(t)\}$ is differentiable at t_0;*

and

(6) *If $f(x,y)$ is differentiable at $\{x(u_0,v_0), y(u_0,v_0)\}$ and $x(u,v)$, $y(u,v)$ are differentiable at (u_0,v_0), then $f\{x(u,v), y(u,v)\}$ is differentiable at (u_0,v_0).*

As an indication of the sort of condition to be looked for, write the condition of simple differentiability in the linear form of chapter IV (7):

(7) *$f(x)$ is differentiable at x_0, if, throughout some neighbourhood*

$$|\delta x| < \eta(x_0),$$

$$\delta f \equiv f(x_0+\delta x)-f(x_0) = A(x_0,\delta x)\,\delta x,$$

where $A(x_0,\delta x)$ converges as $\delta x \to 0$.

The definition of differentiability which we adopt generalizes (7) in the form

(8) *$f(x,y)$ is differentiable† at (x_0,y_0), if, throughout some neighbourhood*

$$|\delta x|, |\delta y| < \eta(x_0,y_0),$$

$$\delta f \equiv f(x_0+\delta x, y_0+\delta y)-f(x_0,y_0)$$
$$= A(x_0,y_0,\delta x,\delta y)\,\delta x + B(x_0,y_0,\delta x,\delta y)\,\delta y,$$

where A, B converge as $\delta x, \delta y \to 0$.

Since A, B converge, they are bounded near $\delta x = 0 = \delta y$, i.e. we have

$$|A|, |B| < \text{some } M(x_0,y_0),$$

if $$|\delta x|, |\delta y| < \text{some } \eta(x_0,y_0),$$

and so $$|\delta f| < (|\delta x|+|\delta y|)M$$
$$< \epsilon, \text{ if } |\delta x|, |\delta y| < \tfrac{1}{2}\epsilon/M.$$

Thus

(9) *$f(x,y)$ is continuous at (x_0,y_0), if it is differentiable there.*

To prove (6), we have, if $|\delta u|, |\delta v| <$ some $\eta_1(u_0,v_0)$,

$$\delta x = A_1\,\delta u + B_1\,\delta v, \qquad \delta y = A_2\,\delta u + B_2\,\delta v,$$

where A_1, B_1, A_2, B_2 are functions of $u_0, v_0, \delta u, \delta v$ which converge as $\delta u, \delta v \to 0$. Hence, by substitution in (8), if $|\delta x|, |\delta y| <$ some $\eta(x_0,y_0)$,

$$\delta f = (A_1 A + A_2 B)\,\delta u + (B_1 A + B_2 B)\,\delta v, \tag{10}$$

where A, B are functions of x, y, x_0, y_0, and therefore of $u, v, \delta u, \delta v$, which converge as $\delta x, \delta y \to 0$. But, by (9), since $x(u,v)$, $y(u,v)$ are differentiable,

$$\delta x, \delta y \to 0 \text{ as } \delta u, \delta v \to 0,$$

† We therefore distinguish between 'differentiable in (x, y)' and 'partially differentiable in x and in y'.

and $\qquad |\delta x|, |\delta y| < \eta(x_0, y_0)$, if $|\delta u|, |\delta v| <$ some $\eta_2(u_0, v_0)$.

Hence (10) holds, if $\qquad |\delta u|, |\delta v| <$ both η_1, η_2;

and $$A_1 A + A_2 B, \qquad B_1 A + B_2 B,$$

being sums of products of convergent functions, converge as $\delta u, \delta v \to 0$. Thus (10) is exactly the condition of differentiability of the function $f\{x(u,v), y(u,v)\}$, and (6) is established.

The simple condition of differentiability (7) comes from the generalized condition (8) by mere suppression of the variation δy, and therefore (5) appears as a particular case of (6).

3. Conditions sufficient for differentiability

It should be remarked that the condition of differentiability along every curve through a point is less stringent than the condition of differentiability (8). For consider the function

$$f(x, y) = |xy|^{\frac{1}{2}} \tag{11}$$

at the origin. If $x(t)$, $y(t)$ are differentiable at t_0, corresponding to $x = 0 = y$, we have near this point

$$\delta x = A(t_0, \delta t)\, \delta t, \qquad \delta y = B(t_0, \delta t)\, \delta t,$$

where A, B converge as $\delta t \to 0$. Hence

$$\delta f = |\delta x \delta y|^{\frac{1}{2}} = |AB|^{\frac{1}{2}}\, \delta t,$$

where $|AB|^{\frac{1}{2}}$ converges as $\delta t \to 0$. Thus $f\{x(t), y(t)\}$ is differentiable at t_0. But the formula (8) requires a relation

$$|\delta x \delta y|^{\frac{1}{2}} = A(\delta x, \delta y)\, \delta x + B(\delta x, \delta y)\, \delta y,$$

where A, B converge, say to A_0, B_0, as $\delta x, \delta y \to 0$. If we pass to the origin successively along the lines $\delta x = 0$, $\delta y = 0$, $\delta x = \delta y$, we are led respectively to the three incompatible relations

$$B_0 = 0, \quad A_0 = 0, \quad A_0 + B_0 = 1.$$

More generally, it can be seen that (5) is sufficiently secured, if δf is expressible as a homogeneous function (not necessarily either rational or integral†) in $\delta x, \delta y$ of the first degree, with coefficients, also involving $\delta x, \delta y$, which converge as $\delta x, \delta y \to 0$. Since homogeneous functions of the first degree of homogeneous functions of the first degree are themselves homogeneous of the first degree, it is clear that such a form for δf also secures (6).

The condition given in (8) goes further in requiring the homogeneous

† But, if δf is fractional in δx, δy, the denominator must not vanish *near* $\delta x = 0 = \delta y$ (though of course it vanishes *at* the point).

function to be rational and integral (i.e. linear). It has, however, the merit of greater simplicity and is generally satisfied by the elementary functions. This follows sufficiently from the theorem:

(12) *$f(x,y)$ is differentiable at a point at which both $f_x(x,y)$, $f_y(x,y)$ exist and one of them is continuous.*

For near such a point we may write

$$\begin{aligned}\delta f &= \{f(x+\delta x, y+\delta y)-f(x, y+\delta y)\}+\{f(x, y+\delta y)-f(x,y)\}\\ &= f_x(x+\theta\,\delta x, y+\delta y)\,\delta x+B(x,y,\delta y)\,\delta y,\end{aligned}$$

by the formula of the mean and by (7), respectively. Hence, if $f_x(x,y)$ is continuous at the point, the conditions of (8) are secured.

The conditions of (12) are of course not necessary to secure (8), as is seen by considering a function of the form

$$f(x,y) \equiv \phi(x)+\psi(y).$$

For then differentiability is secured by mere existence of the derivatives $\phi'(x), \psi'(y)$ without reference to their continuity.

Since the partial derivatives of a polynomial are themselves polynomials, they are everywhere continuous, and so, by (12), the conditions of (8) everywhere apply. Moreover, the elementary functions are, in general, differentiable and their derived functions continuous, save at zeros and infinities. Thus the condition (8) holds at most points for most functions which occur naturally in analysis.

4. Differentiability of $f(x_1,\ldots,x_n)$

The foregoing principles are extended without difficulty to $f(x_1,\ldots,x_n)$, a function of n arguments. The partial derivative in x_1, written, as we may prefer, $\partial f/\partial x$, or $f_1(x_1,\ldots,x_n)$, is obtained by keeping $x_2,\ldots,x_n$ constant and measures variation in x_1 alone.

For $f(x_1,\ldots,x_n)$ the corresponding definition of differentiability is:

(13) *$f(x_1,\ldots,x_n)$ is differentiable at $(x_1,\ldots,x_n)$, if, in the neighbourhood of that point, we can write*

$$\delta f = A_1\,\delta x_1+\ldots+A_n\,\delta x_n,$$

where $A_1,\ldots,A_n$ are functions of $x_1,\ldots,x_n$, $\delta x_1,\ldots,\delta x_n$ which converge as $\delta x_1,\ldots,\delta x_n \to 0$.

We can now state (6) in the more general form:

(14) *If the n arguments of a differentiable function are replaced by differentiable functions of other p arguments ($p \gtreqless n$), the transformed function is differentiable.*†

† i.e. briefly: differentiable functions of differentiable functions are themselves differentiable.

The proof follows the proof of (6). We can similarly prove, in extension of (9), that

(15) *A function differentiable at a point is continuous there.*

Finally, we can prove that

(16) *A function is differentiable at a point at which all the partial derivatives exist and all but one are continuous.*

For we write, as in the proof of (12),

$$\begin{aligned}\delta f(x_1,\dots,x_n) &= f(x_1+\delta x_1, x_2+\delta x_2,\dots, x_n+\delta x_n)-f(x_1, x_2+\delta x_2\dots, x_n+\delta x_n)+\\ &\quad +f(x_1, x_2+\delta x_2,\dots, x_n+\delta x_n)-f(x_1, x_2,\dots, x_n+\delta x_n)+\\ &\quad +\dots+\\ &\quad +f(x_1, x_2,\dots, x_{n-1}, x_n+\delta x_n)-f(x_1, x_2,\dots, x_{n-1}, x_n)\\ &= f_1(x_1+\theta_1\,\delta x_1, x_2+\delta x_2,\dots, x_n+\delta x_n)\,\delta x_1 +\\ &\quad +f_2(x_1, x_2+\theta_2\,\delta x_2,\dots, x_n+\delta x_n)\,\delta x_2 +\\ &\quad +\dots+\\ &\quad +A(x_1, x_2,\dots, x_n, \delta x_n)\,\delta x_n,\end{aligned}$$

where A converges as $\delta x_n \to 0$. Thus the differentiability of f is secured by the continuity of $f_1,\dots,f_{n-1}$ alone. More exactly, we can see that this differentiability is sufficiently secured, if f_1 is continuous in $(x_1,\dots,x_n)$, f_2 continuous in $(x_2,\dots,x_n)$,.... and f_{n-1} continuous in (x_{n-1}, x_n).

The form of the proof suggests that, if we are to rely on the continuity of the partial derivatives, the above conditions are the fewest possible, and this is confirmed by examples. Thus, to see that at least one partial derivative must be fully continuous, consider the example

$$f(x,y,z) = \frac{x^3y^3z^3}{x^8+y^8+z^8}, \qquad f(0,0,0) = 0. \tag{17}$$

Then we find that

$$f_x(0,y,z) = 0, \quad f_x(x,0,z) = 0, \quad f_x(x,y,0) = 0,$$

so that f_x is continuous at the origin in (y,z), in (z,x), and in (x,y). By symmetry f_y, f_z are similarly continuous at the origin.

But if $f(x,y,z)$ were differentiable at the origin, then, taking increments x,y,z, we should have a relation

$$\frac{x^3y^3z^3}{x^8+y^8+z^8} = Ax+By+Cz,$$

where, as $(x,y,z) \to (0,0,0)$, the three functions A, B, C converge, say to A_0, B_0, C_0. Successive passage to the origin along the four paths $y=z=0$, $z=x=0$, $x=y=0$, $x=y=z$ leads to the incompatible results,

$$A_0 = 0, \quad B_0 = 0, \quad C_0 = 0, \quad A_0+B_0+C_0 = \tfrac{1}{3}.$$

Again, if f_x is a partial derivative which is fully continuous, then, further, one of the remaining partial derivatives must be continuous in the remaining variables $(y, z, \ldots)$. For, if $f(x, y, z, \ldots)$ is differentiable, so too is $f(\xi, y, z, \ldots)$ where x has been made constant. Hence, by the above argument, if we are to rely on continuity of partial derivatives, one at least of f_y, $f_z, \ldots$ must be fully continuous in $(y, z, \ldots)$; and so forth.

5. The formula of the total differential

The condition of differentiability (13) leads at once to the fundamental formula of the total differential:

(18) *If the arguments of a differentiable function $f(x_1, \ldots, x_n)$ are differentiable functions of a parameter t, then*

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t} + \ldots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t}.$$

For, from (13), we have, on division by δt,

$$\frac{\delta f}{\delta t} = A_1 \frac{\delta x_1}{\delta t} + \ldots + A_n \frac{\delta x_n}{\delta t}.$$

The limit $\delta t \to 0$ carries with it the limit $\delta x_1, \ldots, \delta x_n \to 0$, since $x_1, \ldots, x_n$, are differentiable and therefore continuous functions of t.

Taking this limit we have

$$\frac{\partial f}{\partial t} = A_{10}\frac{\partial x_1}{\partial t} + \ldots + A_{n0}\frac{\partial x_n}{\partial t}, \tag{19}$$

where $A_{10}, \ldots, A_{n0}$ are the limits to which $A_1, \ldots, A_n$ converge as $\delta x_1, \ldots, \delta x_n \to 0$. But, since there is convergence, we may evaluate A_{10} by proceeding to the limit in any manner we please. In particular, if we travel along the path $\delta x_2 = 0 = \ldots = \delta x_n$, we have in (13)

$$\delta f = A_1 \delta x_1 \qquad (x_2, \ldots, x_n \text{ constant}),$$

i.e.
$$\frac{\partial f}{\partial x_1} = A_{10},$$

and similarly for every A_{r0}. Substitution of these values in (19) gives the desired formula.

The arguments $x_1, \ldots, x_n$ may, of course, involve other parameters than t and the symbols of differentiation in t are therefore rightly written as symbols of partial differentiation.

The formula of the total differential can be written in two special ways which are significant. We may suppress f and obtain the *identity of partial differential operators*

$$\frac{\partial}{\partial t} = \frac{\partial x_1}{\partial t}\frac{\partial}{\partial x_1} + \ldots + \frac{\partial x_n}{\partial t}\frac{\partial}{\partial x_n}. \tag{20}$$

This is a fundamental formula in the theory of the transformation of partial differential operators. Again we may suppress t and write

$$Df = \frac{\partial f}{\partial x_1} Dx_1 + \ldots + \frac{\partial f}{\partial x_n} Dx_n, \tag{21}$$

where D denotes *differentiation in an unspecified parameter.*

In the formula (13)

$$\delta f = A_1\,\delta x_1 + \ldots + A_n\,\delta x_n,$$

the coefficients $A_1,\ldots, A_n$ are essentially convergent, as $\delta x_1\ldots, \delta x_n \to 0$, and so, to a prescribed standard of accuracy, may be replaced by their limiting values $\partial f/\partial x_1,\ldots, \partial f/\partial x_n$, if $\delta x_1,\ldots, \delta x_n$ are suitably small. We are thus led to the *approximate* formula

$$\delta f \doteqdot \frac{\partial f}{\partial x_1}\delta x_1 + \ldots + \frac{\partial f}{\partial x_n}\delta x_n,$$

where $\delta x_1,\ldots, \delta x_n$ are small. Evidently, to this degree of approximation, the arguments of f may receive their increments independently and in any order. This corresponds analytically to the principle of Mathematical Physics known as the *independence of small increments.* Thus functions which are differentiable in the extended sense conform to this principle, which, in general, characterizes the functions of Mathematical Physics.

In somewhat similar fashion, in the formula of the total differential we may suppose each argument *in turn* to become the appropriate function of t, differentiate it and add the results in any order. In this form the principle becomes that of *the independence of differentiations,* and is exact. We have already met an instance of this in the formula for the ordinary differentiation of a product

$$D(u_1\ldots u_n) = \sum u_1\ldots u_{r-1}\,u_{r+1}\ldots u_n\,Du_r,$$

It is, in fact, merely the formula of the total differential applied to the particular function $u_1\ldots u_n$.

The following converse of (18) is often needed:

(22) *If, for all functions $x_1(t),\ldots, x_n(t)$ differentiable at t_0, the formula*

$$\frac{\partial f}{\partial t} = \phi_1(x_1,\ldots, x_n)\frac{\partial x_1}{\partial t} + \ldots + \phi_n(x_1,\ldots, x_n)\frac{\partial x_n}{\partial t}$$

holds at t_0, then necessarily

$$\phi_1 = \frac{\partial f}{\partial x_1}, \quad \ldots, \quad \phi_n = \frac{\partial f}{\partial x_2}.$$

If we consider the set of functions

$$x_1 = t, \quad x_2 = \text{constant}, \ \ldots, \ x_n = \text{constant},$$

the given formula at once reduces to

$$\frac{\partial f(x_1,\ldots,x_n)}{\partial x_1} = \phi_1(x_1,\ldots,x_n),$$

and similarly for $\phi_2,\ldots,\phi_n$.

It is often convenient to use the notation of (21) and to say that, if identically

$$Df = \phi_1 Dx_1 + \ldots + \phi_n Dx_n,$$

then $\phi_1 = \partial f/\partial x_1$, etc.

6. Higher partial derivatives

If the partial derivatives of a function $f(x_1,\ldots,x_n)$ exist throughout some region of values of $(x_1,\ldots,x_n)$, the aggregates of their values define n new functions $f_1(x_1,\ldots,x_n),\ldots,f_n(x_1,\ldots,x_n)$, which can become, in their turn, subject-matter for partial differentiation. We are led in this way to the partial derivatives of the second order and ultimately to those of any order.

For simplicity consider $f(x,y)$ a function of two variables. From the partial derivatives of the first order

$$f_x(x,y), \quad f_y(x,y)$$

we obtain the four partial derivatives of the second order

$$\frac{\partial}{\partial x}(f_x), \quad \frac{\partial}{\partial y}(f_x), \quad \frac{\partial}{\partial x}(f_y), \quad \frac{\partial}{\partial y}(f_y),$$

which we write as

$$\frac{\partial^2 f}{\partial x^2}, \quad \frac{\partial^2 f}{\partial y \partial x}, \quad \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial y^2},$$

or as

$$f_{xx}, \quad f_{yx}, \quad f_{xy}, \quad f_{yy}.$$

The differentiations themselves are denoted by the symbols of operation

$$\left(\frac{\partial}{\partial x}\right)^2, \quad \frac{\partial}{\partial y}\frac{\partial}{\partial x}, \quad \frac{\partial}{\partial x}\frac{\partial}{\partial y}, \quad \left(\frac{\partial}{\partial y}\right)^2,$$

or by

$$\partial_x^2, \quad \partial_y \partial_x, \quad \partial_x \partial_y, \quad \partial_y^2.$$

In these symbols, it should be noted, it is always the operation on the right which is first performed. In the case of f_{xy}, f_{yx} this violates the usual convention that that operation is first performed whose symbol comes nearest to the symbol of the operand. Some such violation is necessary, if f_{xy} and $\partial_x \partial_y f$ are to mean the same thing. Happily no confusion is likely to arise in practice, since under certain wide conditions, as we shall presently see, the two operations $\partial_x \partial_y$, $\partial_y \partial_x$ are equivalent. The second derivatives f_{xx}, f_{yy} are merely the second derivatives of $f(x,\eta), f(\xi,y)$, functions of one variable, and so bring in no new principle.

Of the third order there will evidently be eight possible derivatives

$$f_{xxx},\quad f_{xxy},\quad f_{xyx},\quad f_{xyy},\quad f_{yxx},\quad f_{yxy},\quad f_{yyx},\quad f_{yyy},$$

and generally of order r the number of possible derivatives will be 2^r, the number of arrangements of r letters which can be formed out of x, y and their repetitions. So, more generally, in a field of n variables there are n^r possible derivatives of order r.

If, however, the conditions are satisfied under which the operations ∂x, ∂y are interchangeable, these numbers are appreciably reduced. The order of the letters x, y is immaterial, and thus a typical derivative of order r may be written

$$\frac{\partial^r f}{\partial x^s \partial y^{r-s}}.$$

The number of derivatives of order r is now $r+1$ and the number of order *not exceeding* r is $\frac{1}{2}r(r+3)$.

Similarly, in the field of n variables, the typical derivative of order r is

$$\frac{\partial^r f}{\partial x_1^{\alpha_1}\partial x_2^{\alpha_2}\dots\partial x_n^{\alpha_n}},$$

where $\alpha_1+\alpha_2+\dots+\alpha_n = r$.

The number of derivatives of order r is the number of distinct homogeneous polynomials of order r which can be formed from n letters, i.e. it is

$$\frac{(n+r-1)!}{(n-1)!\,r!},$$

while the total number of derivatives of order *not exceeding* r is

$$\frac{(n+r)!}{n!\,r!}-1,$$

or, if for convenience we include the function itself, we may say that the number of such derivatives is

$$\frac{(n+r)!}{n!\,r!}.$$

7. Conditions for the equivalence of $\dfrac{\partial^2 f}{\partial x \partial y}$, $\dfrac{\partial^2 f}{\partial y \partial x}$.

We suppose that both $f_{xy}(\xi,\eta)$ and $f_{yx}(\xi,\eta)$ exist, and therefore that $f_x(x,y)$, $f_y(x,y)$ exist throughout a neighbourhood of (ξ,η). Write

$$\phi(h,k) \equiv \frac{f(\xi+h,\eta+k)-f(\xi+h,\eta)-f(\xi,\eta+k)+f(\xi,\eta)}{hk}. \tag{23}$$

If we write also

$$\psi(x) \equiv \frac{f(x,\eta+k)-f(x,\eta)}{k}, \tag{24}$$

we have $$\phi(h,k) = \frac{\psi(\xi+h)-\psi(\xi)}{h}. \tag{25}$$

Hence $$\lim_{h\to 0}\phi(h,k) = \psi'(\xi) = \frac{f_x(\xi,\eta+k)-f_x(\xi,\eta)}{k},$$

and therefore also $$\lim_{k\to 0}\lim_{h\to 0}\phi(h,k) = f_{yx}(\xi,\eta).$$

It follows that

(26) *The partial derivatives* f_{xy}, f_{yx} *are equivalent, if and only if the repeated limits*
$$\lim_{k\to 0}\lim_{h\to 0}\phi(h,k), \qquad \lim_{h\to 0}\lim_{k\to 0}\phi(h,k)$$
are the same.

The equivalence of the repeated limits is thus sufficiently secured by the continuity of $\phi(h,k)$ at $(0,0)$. We can accordingly show that

(27) *The partial derivatives f_{xy}, f_{yx} are equal at (ξ,η), if one of them exists near (ξ,η) and is continuous there.*

For, applying the formula of the mean to (25), we have
$$\begin{aligned}\phi(h,k) &= \psi'(\xi+\theta h) \qquad (0<\theta<1)\\ &= \frac{f_x(\xi+\theta h,\eta+k)-f_x(\xi+\theta h,\eta)}{k}, \text{ from (24)}, \qquad (28)\\ &= f_{yx}(\xi+\theta h,\eta+\theta' k) \qquad (0<\theta'<1),\end{aligned}$$
by a second application of the formula of the mean, on the supposition that $f_{yx}(x,y)$ exists near (ξ,η). Hence, if $f_{yx}(x,y)$ is also continuous at (ξ,η), then $\phi(h,k)$ is continuous at $(0,0)$ and $f_{xy}(\xi,\eta)$, $f_{yx}(\xi,\eta)$ are the same.

The conditions of (27) evidently cover the case of polynomials and, in general, of elementary functions at most points. Thus we expect that for operation on the elementary functions the operators ∂_x, ∂_y are, in general, commutative except possibly at isolated points.

Conditions sufficient for the equivalence of f_{xy}, f_{yx} can also be given in terms of the differentiability of $f_x(x,y), f_y(x,y)$. We have the theorem due to Young:

(29) *If $f_x(x,y)$, $f_y(x,y)$ are both differentiable at (ξ,η), then $f_{xy}(\xi,\eta)$, $f_{yx}(\xi,\eta)$ both exist and are equal.*

For suppose that, in some neighbourhood of (ξ,η),
$$f_x(\xi+h,\eta+k)-f_x(\xi,\eta) = A(\xi,\eta,h,k)h+B(\xi,\eta,h,k)k, \tag{30}$$
where A, B converge as $h,k\to 0$. By consideration of the special paths $k=0$, $h=0$ we know that they must converge respectively to $f_{xx}(\xi,\eta)$, $f_{yx}(\xi,\eta)$ and that accordingly these partial derivatives must exist.

By putting $k = h$ in (28) we have

$$\phi(h,h) = \frac{f_x(\xi+\theta h, \eta+h)-f_x(\xi+\theta h, \eta)}{h}$$

$$= \theta\{A(\xi,\eta,\theta h,h)-A(\xi,\eta,\theta h,0)\}+B(\xi,\eta,\theta h,h),$$

on substitution from (30). As $h \to 0$, the difference of the A's vanishes, since A is convergent, and B converges to $f_{yx}(\xi,\eta)$.

Hence, if $f_x(x,y)$ is differentiable at (ξ,η),

$$\lim_{h\to 0}\phi(h,h) = f_{yx}(\xi,\eta).$$

If $f_y(x,y)$ is also differentiable at (ξ,η), the limit of $\phi(h,h)$ is similarly equal to $f_{xy}(\xi,\eta)$ and the existence and equality of $f_{xy}(\xi,\eta)$, $f_{yx}(\xi,\eta)$ are therefore established.

We may observe that of the two conditions (27), (29), which independently secure the equality of $f_{xy}(\xi,\eta)$, $f_{yx}(\xi,\eta)$, neither includes the other. Decisive examples are given by

$$f(x,y) \equiv F(x)+G(y), \tag{31}$$

$$f(x,y) \equiv F(x+y). \tag{32}$$

In (31), $f_{xy}(x,y)$ is everywhere zero and so everywhere continuous, so long as $F'(x)$, $G'(y)$ exist. But $f_x(x,y)$, $f_y(x,y)$ are differentiable, only if $F'(x)$, $G'(y)$ are also differentiable. In (32), $f_x(x,y)$, $f_y(x,y)$ are differentiable at (ξ,η), if $F''(x+y)$ exists at $\xi+\eta$. But $f_{xy}(x,y)$, $f_{yx}(x,y)$ are continuous at (ξ,η), only if $F''(x+y)$ is continuous at $\xi+\eta$.

8. Generalization of the formula of the mean

We have already, in chapter V, obtained Taylor's formula as a generalization of the formula of the mean to the case of n differentiations. We now generalize it for the field of n variables.

Suppose in the field of n variables $(x_1,\ldots,x_n)$ that there is an interval

$$a_r \leqslant x_r \leqslant a_r+h_r \qquad (r = 1,\ldots,n)$$

throughout which $f(x_1,\ldots,x_n)$ is differentiable. Write

$$F(t) = f(a_1+h_1 t,\ldots,a_n+h_n t),$$

so that $\qquad F(1)-F(0) = f(a_1+h_1,\ldots,a_n+h_n)-f(a_1,\ldots,a_n).$

Then $F(t)$ is differentiable throughout the interval $(0, 1)$ and so, by the theorem of the mean,

$$F(1)-F(0) = F'(\theta) \qquad (0<\theta<1).$$

By the formula of the total differential,

$$F'(t) = h_1 f_1(a_1+h_1 t,\ldots,a_n+h_n t)+\ldots+h_n f_n(a_1+h_1 t,\ldots,a_n+h_n t).$$

Hence we have the *generalized linear formula of the mean*.

(33) $f(a_1+h_1, \ldots, a_n+h_n)-f(a_1, \ldots, a_n)$
$$= h_1 f_1(a_1+\theta h_1, \ldots, a_n+\theta h_n)+\ldots+h_n f_n(a_1+\theta h_1, \ldots, a_n+\theta h_n),$$
for some θ *in* $)0, 1($.

Given a second function $g(x_1, \ldots, x_n)$ differentiable throughout the same interval we similarly define
$$G(t) = g(a_1+h_1 t, \ldots, a_n+h_n t),$$
and so from the fractional formula of the mean obtain *the generalized fractional formula of the mean*

(34) $$\frac{f(a_1+h_1, \ldots, a_n+h_n)-f(a_1, \ldots, a_n)}{g(a_1+h_1, \ldots, a_n+h_n)-g(a_1, \ldots, a_n)}$$
$$= \frac{h_1 f_1(a_1+\theta h_1, \ldots, a_n+\theta h_n)+\ldots+h_n f_n(a_1+\theta h_1, \ldots, a_n+\theta h_n)}{h_1 g_1(a_1+\theta h_1, \ldots, a_n+\theta h_n)+\ldots+h_n g_n(a_1+\theta h_1, \ldots, a_n+\theta h_n)},$$
for some θ *in* $)0, 1($, *provided, of course, that the second fraction cannot take the form* $0/0$ *in the interval considered*.

On the other hand we may write
$$F(t) = f(a_1+h_1 t, a_2+h_2, \ldots, a_n+h_n)+f(a_1, a_2+h_2 t, \ldots, a_n+h_n)+$$
$$+\ldots+f(a_1, a_2, \ldots, a_n+h_n t),$$
where the arguments of f on its rth appearance are
$$a_1, \ldots, a_{r-1}, a_r+h_r t, a_{r+1}+h_{r+1}, \ldots, a_n+h_n.$$
We still have
$$F(1)-F(0) = f(a_1+h_1, \ldots, a_n+h_n)-f(a_1, \ldots, a_n),$$
but now
$$F'(t) = h_1 f_1(a_1+h_1 t, a_2+h_2, \ldots, a_n+h_n)+\ldots+h_n f_n(a_1, a_2, \ldots, a_n+h_n t),$$
and the corresponding generalized linear formula of the mean is

(35) $f(a_1+h_1, \ldots, a_n+h_n)-f(a_1, \ldots, a_n)$
$$= h_1 f_1(a_1+\theta h_1, a_2+h_2, \ldots, a_n+h_n)+h_2 f_2(a_1, a_2+\theta h_2, \ldots, a_n+h_n)+$$
$$+\ldots+h_n f_n(a_1, a_2, \ldots, a_n+\theta h_n),$$
for some θ *in* $)0, 1($.

There is, of course, a corresponding generalized fractional formula.

We may evidently devise other forms for $F(t)$, in a sense intermediate between those already given, which preserve the identity
$$F(1)-F(0) = f(a_1+h_1, \ldots, a_n+h_n)-f(a_1, \ldots, a_n).$$
We have essentially to distribute the n arguments $a_r+h_r t$ over the sum of not more than n forms f, filling up the remaining arguments with a_r or

a_r+h_r by a rule which is easier to see than to state. Thus, for instance, we may write

$$F(t) \equiv f(a_1+h_1 t,\dots, a_{n-1}+h_{n-1} t, a_n+h_n)+f(a_1,\dots, a_{n-1}, a_n+h_n t)$$

or

$$F(t) \equiv f(a_1+h_1 t, a_2+h_2 t, a_3,\dots, a_n)+f(a_1+h_1, a_2+h_2, a_3+h_3 t,\dots, a_n+h_n t).$$

Each such $F(t)$ evidently gives rise to a distinct generalization of the linear formula of the mean. In the fractional formula, for even greater generality, we may associate one form for $F(t)$ with another form for $G(t)$.

9. Generalization of Taylor's formula

We may similarly generalize Taylor's formula for the field of n variables. For suppose that $f(x_1,\dots,x_n)$ and its partial derivatives of the first $p-1$ orders are differentiable throughout the 'interval'

$$a_r \leqslant x_r \leqslant a_r+h_r \qquad (r=1,\dots,n).$$

As before, we write

$$F(t) \equiv f(a_1+h_1 t,\dots, a_n+h_n t).$$

It follows that $F(t)$ has derivatives of the first p orders throughout $(0, 1)$. Precisely, by the formula of the total differential,

$$\begin{aligned} F'(t) &= h_1 f_1(a_1+h_1 t,\dots, a_n+h_n t)+\dots+h_n f_n(a_1+h_1 t,\dots, a_n+h_n t) \\ &= \left(h_1 \frac{\partial}{\partial a_1}+\dots+h_n \frac{\partial}{\partial a_n}\right) f(a_1+h_1 t,\dots, a_n+h_n t) \\ &\equiv \phi(a_1+h_1 t,\dots, a_n+h_n t), \text{ say.} \end{aligned}$$

Since the first derivatives of $f(x_1,\dots,x_n)$ are all differentiable throughout the stated interval, $\phi(x_1,\dots,x_n)$ is likewise differentiable throughout the interval and therefore, by a second application of the formula of the total differential,

$$\begin{aligned} F''(t) &= \left(h_1 \frac{\partial}{\partial a_1}+\dots+h_n \frac{\partial}{\partial a_n}\right) \phi(a_1+h_1 t,\dots, a_n+h_n t) \\ &= \left(h_1 \frac{\partial}{\partial a_1}+\dots+h_n \frac{\partial}{\partial a_n}\right)\left(h_1 \frac{\partial}{\partial a_1}+\dots+h_n \frac{\partial}{\partial a_n}\right) f(a_1+h_1 t,\dots, a_n+h_n t), \end{aligned}$$

which we may conveniently write

$$F''(t) = \left(h_1 \frac{\partial}{\partial a_1}+\dots+h_n \frac{\partial}{\partial a_n}\right)^2 f(a_1+h_1 t,\dots, a_n+h_n t).$$

Proceeding in this way we have at length for every $r \leqslant p$ and every t in $(0, 1)$

$$F_r(t) = \left(h_1 \frac{\partial}{\partial a_1}+\dots+h_n \frac{\partial}{\partial a_n}\right)^r f(a_1+h_1 t,\dots, a_n+h_n t).$$

If we write the Taylor's series for $F(t)$ with Lagrange's form of the remainder,† we have

$$F(1)-\sum_{r=0}^{p-1}\frac{F_r(0)}{r!}=\frac{F_p(\theta)}{p!}\qquad (0<\theta<1).$$

On substitution for F this gives

$$(36)\quad f(a_1+h_1,\ldots,a_n+h_n)-\sum_{r=0}^{p-1}\frac{1}{r!}\left(h_1\frac{\partial}{\partial a_1}+\ldots+h_n\frac{\partial}{\partial a_n}\right)^r f(a_1,\ldots,a_n)$$
$$=\frac{1}{p!}\left(h_1\frac{\partial}{\partial a_1}+\ldots+h_n\frac{\partial}{\partial a_n}\right)^p f(a_1+\theta h_1,\ldots,a_n+\theta h_n),$$

for some θ in $)0,1($.

We may evidently derive corresponding generalizations from other forms of the remainder in Taylor's series. Similarly from the fractional form‡ we obtain

$$(37)\quad \frac{f(a_1+h_1,\ldots,a_n+h_n)-\sum_{r=0}^{p-1}\frac{1}{r!}\left(h_1\frac{\partial}{\partial a_1}+\ldots+h_n\frac{\partial}{\partial a_n}\right)^r f(a_1,\ldots,a_n)}{g(a_1+h_1,\ldots,a_n+h_n)-\sum_{r=0}^{p-1}\frac{1}{r!}\left(h_1\frac{\partial}{\partial a_1}+\ldots+h_n\frac{\partial}{\partial a_n}\right)^r g(a_1,\ldots,a_n)}$$
$$=\frac{\left(h_1\frac{\partial}{\partial a_1}+\ldots+h_n\frac{\partial}{\partial a_n}\right)^p f(a_1+\theta h_1,\ldots,a_n+\theta h_n)}{\left(h_1\frac{\partial}{\partial a_1}+\ldots+h_n\frac{\partial}{\partial a_n}\right)^p g(a_1+\theta h_1,\ldots,a_n+\theta h_n)},$$

for some θ in $)0,1($.

We may, of course, obtain other formulae analogous to (36) and (37) by adopting some of the alternative forms that have been indicated for $F(t)$. In point of actual practice such other forms are more curious than useful, and (36) remains the standard expression of Taylor's formula in a field of n variables. Differential operators of the form

$$h_1\frac{\partial}{\partial x_1}+h_2\frac{\partial}{\partial x_2}+\ldots+h_n\frac{\partial}{\partial x_n}$$

in which the coefficients h of the individual operators $\partial/\partial x$ are unaffected by these differentiations are of frequent occurrence in analysis and may be called *Taylor's operators.*

We can generalize for functions of many variables the theorem of

† Chapter V (28), (29). ‡ Chapter V (27).

Taylor's limit.† It is sufficient to give the theorem for two variables in the form:

(38) *If $f(x,y)$ has partial derivatives of order $n-1$ near (a,b) and if at (a,b) these derivatives are themselves differentiable, then near (a,b) we may write*

$$f(a+h,b+k) = \sum_{r=0}^{n} \frac{1}{r!}\left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}\right)^r f(a,b)+\sum_{s=0}^{n} \frac{h^{n-s}k^s}{(n-s)!\,s!}A_{n-s},$$

where every A_{n-s} is a function of a, b, h, k which converges to zero as $h, k \to 0$.

In the proof take $n = 2$ for brevity. Write

$$F(h,k) \equiv f(a+h,b+k)-\frac{1}{2!}\{h^2 f_{11}(a,b)+2hkf_{12}(a,b)+k^2 f_{22}(a,b)\}, \quad (39)$$

so that $F(h,k)$ is differentiable near $h, k = 0$. Then, by the formula of the mean, (33) above, we may write

$$F(h,k) = F(0,0)+h\,F_1(\theta h,\theta k)+k\,F_2(\theta h,\theta k). \quad (40)$$

But, from (39),

$$\begin{aligned} F_1(h,k) &= f_1(a+h,b+k)-hf_{11}(a,b)-kf_{12}(a,b) \\ &= f_1(a,b)+h\epsilon_1+k\eta_1, \end{aligned}$$

where $\epsilon_1, \eta_1 \to 0$ as $h, k \to 0$, since f_1 is differentiable near (a,b). Substituting this and the corresponding expression for $F_2(h,k)$ in (40), we have

$$F(h,k) = F(0,0)+hf_1(a,b)+kf_2(a,b)+h^2\theta\epsilon_1+hk\,\theta(\eta_1+\epsilon_2)+k^2\theta\eta_2,$$

i.e., on substitution for F,

$$\begin{aligned} f(a+h,b+k) = f(a,b)&+hf_1(a,b)+kf_2(a,b)+ \\ &+\frac{1}{2!}\{h^2 f_{11}(a,b)+2hkf_{12}(a,b)+k^2 f_{22}(a,b)\}+ \\ &+h^2\theta\epsilon_1+hk\,\theta(\eta_1+\epsilon_2)+k^2\theta\eta_2, \end{aligned}$$

where $\theta\epsilon_1$, $\theta(\eta_1+\epsilon_2)$, $\theta\eta_2 \to 0$ as $h, k \to 0$. This proves the theorem for the case $n = 2$. The proof for the general case proceeds along similar lines.

We can also generalize for functions of many variables the theorems (38), (43) of chapter V, which express the nth derivative of a function

† Chapter V (32).

of a function in terms of the derivatives of the component functions. The theorems, given for brevity for functions of two variables only, are

(41) *If $f(x,y)$ is n times differentiable and if $D^n x$, $D^n y$ exist, then also $D^n f(x,y)$ exists and is the coefficient of h^n in*

$$n!\sum_{r=0}^{n}\left\{\left(hDx+\ldots+\frac{h^n}{n!}D^n x\right)\frac{\partial}{\partial x}+\left(hDy+\ldots+\frac{h^n}{n!}D^n y\right)\frac{\partial}{\partial y}\right\}^r\frac{f(x,y)}{r!},$$

where the operators $\partial/\partial x$, $\partial/\partial y$ operate only on $f(x,y)$;

and the similar theorem for umbral derivatives:

(42) *If $f(x,y)$ is n times differentiable and if the umbral derivatives $[D^n x]$, $[D^n y]$ exist at some point, then the umbral derivative $[D^n f(x,y)]$ also exists at the point and is the coefficient of h^n in*

$$n!\sum_{r=0}^{n}\left\{\left(h\,Dx+\ldots+\frac{h^n}{n!}[D^n x]\frac{\partial}{\partial x}\right)+\left(h\,Dy+\ldots+\frac{h^n}{n!}[D^n y]\right)\frac{\partial}{\partial y}\right\}^r\frac{f(x,y)}{r!},$$

where the operators $\partial/\partial x$, $\partial/\partial y$ operate only on $f(x,y)$.

In (41) the existence of the derivative can be shown without difficulty by repeated use of the formula of the total differential; the formula for the derivative is sufficiently covered by theorem (42), since, if the full derivative of any order exists, the corresponding umbral derivative also exists and is identical with it. It is therefore sufficient to give a proof of (42) only. This will follow the lines of chapter V (43), the analogous theorem for the single variable.

Let x,y become X, Y, when the fundamental parameter increases by h. Then, since $[D^n x]$, $[D^n y]$ exist,

$$X-x=h\,Dx+\frac{h^2}{2!}[D^2 x]+\ldots+\frac{h^n}{n!}([D^n x]+\epsilon),$$

$$Y-y=h\,Dy+\frac{h^2}{2!}[D^2 y]+\ldots+\frac{h^n}{n!}([D^n y]+\epsilon'),$$

where $\epsilon,\epsilon'\to 0$ as $h\to 0$.

By (38), since f is n times differentiable,

$$f(X,Y)=\sum_{r=0}^{n}\left\{(X-x)\frac{\partial}{\partial x}+(Y-y)\frac{\partial}{\partial y}\right\}^r\frac{f(x,y)}{r!}+\sum_{s=0}^{n}\eta_s(X-x)^{n-s}(Y-y)^s,$$

where the operators $\partial/\partial x$, $\partial/\partial y$ operate only on f and where every $\eta_s\to 0$ as $X-x$, $Y-y\to 0$. Thus

$$f(X,Y)=\sum_{r=0}^{n}\left\{\left(h\,Dx+\ldots+\frac{h^n}{n!}[D^n x]+\frac{h^n\epsilon}{n!}\right)\frac{\partial}{\partial x}+\right.$$

$$\left.+\left(h\,Dy+\ldots+\frac{h^n}{n!}[D^n y]+\frac{h^n\epsilon'}{n!}\right)\frac{\partial}{\partial y}\right\}^r\frac{f(x,y)}{r!}+\sum_{s=0}^{n}\eta_s(X-x)^{n-s}(Y-y)^s. \quad (43)$$

Now assume the theorem (42) to be true for values of the index up to $n-1$ inclusive. Then we have that

$$\sum_{r=0}^{n-1} \frac{h^r}{r!}[D^r f(x,y)]$$

consists of the terms up to and including h^{n-1} in the expansion, in powers of h, of

$$\sum_{r=0}^{n-1} \left\{\left(h\,Dx+\ldots+\frac{h^{n-1}}{(n-1)!}[D^{n-1}x]\right)\frac{\partial}{\partial x}+\right.$$
$$\left.+\left(h\,Dy+\ldots+\frac{h^{n-1}}{(n-1)!}[D^{n-1}y]\right)\frac{\partial}{\partial y}\right\}^r \frac{f(x,y)}{r!}.$$

Hence, by (43),
$$\frac{f(X,Y)-\sum_{r=0}^{n-1}\frac{h^r}{r!}[D^r f(x,y)]}{h^n/n!} \tag{44}$$

contains no negative powers of h. Also ϵ, ϵ', η_s $(s = 0,\ldots,n)$ all converge to zero as $h \to 0$. Thus, if the theorem holds for indices up to $n-1$ inclusive, $[D^n f]$ exists, and, moreover, to evaluate $[D^n f]$ it is enough to retain the coefficient of $h^n n!$ in the numerator of (44), i.e. in $f(X,Y)$.

In other words, $[D^n f]$ is the coefficient of h^n in

$$n!\sum_{r=0}^{n}\left\{\left(h\,Dx+\ldots+\frac{h^n}{n!}[D^n x]\right)\frac{\partial}{\partial x}+\left(h\,Dy+\ldots+\frac{h^n}{n!}[D^n y]\right)\frac{\partial}{\partial y}\right\}^r\frac{f(x,y)}{r!}.$$

Thus the theorem, if true for indices up to and including $n-1$, is true also for the index n. It is evidently true for $n = 1$, and the induction is therefore established.

In (41) write for clearness

$$x_r,\ y_r = \frac{D^r x}{r!},\ \frac{D^r y}{r!}$$

and, regarding all the variables $x, y, x_1, y_1, \ldots, x_n, y_n$ as independent, differentiate partially in x_p where $p \leqslant n$. Then we have that

$$\frac{\partial}{\partial x_p}\left\{\frac{D^n f(x,y)}{n!}\right\}$$

is the coefficient of h^n in

$$\sum_{r=1}^{n}\left\{(hx_1+\ldots+h^n x_n)\frac{\partial}{\partial x}+(hy_1+\ldots+h^n y_n)\frac{\partial}{\partial y}\right\}^{r-1} h^p \frac{\partial}{\partial x}\frac{f(x,y)}{(r-1)!},$$

i.e. is the coefficient of h^{n-p} in

$$\sum_{s=1}^{n}\left\{(hx_1+\ldots+h^n x_n)\frac{\partial}{\partial x}+(hy_1+\ldots+h^n y_n)\frac{\partial}{\partial y}\right\}^s\frac{f_x(x,y)}{s!},$$

where $s \equiv r-1$. But, by (41), this coefficient of h^{n-p} is exactly $D^{n-p}f_x(x,y)/(n-p)!$. We therefore have the following corollary, of which we shall make much use in discussing certain problems of maxima and minima:

(45) *If* $x_r, y_r \equiv \dfrac{D^r x}{r!}, \dfrac{D^r y}{r!}$ *and if* $f(x,y)$ *be* n *times differentiable, so that we can write* $D^n f(x,y)$ *as a function of* $x, y, x_1, y_1,\ldots, x_n, y_n$, *then*

$$\frac{\partial}{\partial x_r}\left\{\frac{D^n f(x,y)}{n!}\right\} = \frac{D^{n-r}f_x(x,y)}{(n-r)!}$$

and so for y.

A similar corollary for umbral derivatives is evidently to be obtained from (42).

10. Euler's theorem on homogeneous functions

By a *homogeneous function* we mean, of course, not merely a polynomial in which the sum of the indices of every term is the same, but more generally any function $f(x_1,\ldots,x_n)$ which satisfies an identity

$$f(x_1 t,\ldots,x_n t) = t^p f(x_1,\ldots,x_n), \tag{46}$$

where p is some constant, most frequently a positive integer. The function $f(x_1,\ldots,x_n)$ is then said to be 'homogeneous of degree p in $x_1,\ldots,x_n$'. For such a function we can prove Euler's theorem:

(47) *A function* $f(x_1,\ldots,x_n)$ *homogeneous of degree* p *in* $x_1,\ldots,x_n$ *satisfies the relation*

$$x_1\frac{\partial f}{\partial x_1}+\ldots+x_n\frac{\partial f}{\partial x_n} = pf.$$

Write
$$X_1 = x_1 t,\ldots,\quad X_n = x_n t \tag{48}$$

so that the defining identity (46) may be written

$$f(X_1,\ldots,X_n) = t^p f(x_1,\ldots,x_n). \tag{49}$$

For differentiation in t, regard $x_1,\ldots,x_n$ as constants. For operation on $f(X_1,\ldots,X_n)$, we then have the identity of operators

$$\frac{\partial}{\partial t} = x_1\frac{\partial}{\partial X_1}+\ldots+x_n\frac{\partial}{\partial X_n},$$

i.e.
$$t\frac{\partial}{\partial t} = X_1\frac{\partial}{\partial X_1}+\ldots+X_n\frac{\partial}{\partial X_n}. \tag{50}$$

On (49) this gives

$$\left(X_1\frac{\partial}{\partial X_1}+...+X_n\frac{\partial}{\partial X_n}\right)f(X_1,..., X_n)$$
$$= t\frac{\partial}{\partial t}\{t^p f(x_1,..., x_n)\}$$
$$= pt^p f(x_1,..., x_n), \text{ since the } x\text{'s are constants as regards } t,$$
$$= pf(X_1,..., X_n).$$

This is exactly Euler's formula stated in terms of arguments X_r. We may equally well replace them by arguments x_r and so (47) is established.

Conversely, we can prove that

(51) *Every function $f(x_1,..., x_n)$ which satisfies the relation*

$$x_1\frac{\partial f}{\partial x_1}+...+x_n\frac{\partial f}{\partial x_n} = pf$$

is a homogeneous function of degree p.

For, with the same substitution (48), we have the same identity of operators (50), and so $f(X_1,..., X_n)$, since it satisfies the relation

$$X_1\frac{\partial f}{\partial X_1}+...+X_n\frac{\partial f}{\partial X_n} = pf,$$

satisfies also the relation $\qquad t\frac{\partial f}{\partial t} = pf,$

which we may equally well write

$$t\frac{df}{dt} = pf,$$

since only one independent variable t is involved. The solution of this is known to be

$$f \equiv f(X_1,..., X_n) = Ct^p, \tag{52}$$

where C is a constant as regards t; that is to say, we may give t any value without altering the value of C. In particular put $t = 1$. Then every X_r becomes x_r and (52) becomes

$$f(x_1,..., x_n) = C.$$

Elimination of C gives

$$f(x_1 t,..., x_n t) = t^p f(x_1,..., x_n),$$

the defining equation of a function homogeneous of order p.

The relation (47) satisfied by the homogeneous function, since it is stated in terms of first-order partial derivatives, is known as a 'partial differential equation of the first order': in particular, it is often known as 'Euler's differential equation for homogeneous functions'.

We may generalize the idea of *homogeneous function* into that of *isobaric function*. Every argument x_r has now its assigned dimension or *weight* p_r. The defining equation of the isobaric function $f(x_1,..., x_n)$ of weight p is now

$$f(x_1 t^{p_1},..., x_n t^{p_n}) = t^p f(x_1,..., x_n). \tag{53}$$

Such an isobaric function arises naturally, if in a set of arguments $u_1,..., u_m$ we have the n functions $f_r(u_1,..., u_m)$ homogeneous of respective degrees p_r. For then a function $F(f_1,...,f_n)$ which is homogeneous in $u_1,..., u_m$ is necessarily isobaric in $f_1,...,f_n$.

It is often sufficient to suppose that the argument x_r is of the weight r indicated by its suffix. We can then prove that

(54) *The isobaric function $f(x_0, x_1, x_2,..., x_n)$ of weight p satisfies the partial differential equation*

$$x_1\frac{\partial f}{\partial x_1}+2x_2\frac{\partial f}{\partial x_2}+...+nx_n\frac{\partial f}{\partial x_n} = pf.$$

Conversely, every solution of the differential equation is such an isobaric function.

We now employ the substitution

$$X_r = x_r t^r \qquad (r = 0, 1, 2,..., n). \tag{55}$$

Again, regarding the x_r as constants for differentiation in t, we have the identity of operators

$$\frac{\partial}{\partial t} = \sum_{r=0}^{n} r x_r t^{r-1}\frac{\partial}{\partial X_r},$$

i.e.

$$t\frac{\partial}{\partial t} = \sum_{r=0}^{n} rX_r\frac{\partial}{\partial X_r}. \tag{56}$$

Operation on the isobaric identity

$$f(X_0, X_1,..., X_n) = t^p f(x_0, x_1,..., x_n) \tag{57}$$

at once gives that

$$X_1\frac{\partial f}{\partial X_1}+2X_2\frac{\partial f}{\partial X_2}+...+nX_n\frac{\partial f}{\partial X_n} = pf, \tag{58}$$

which is the differential equation required.

Conversely, if we retain the substitution (55), we retain also the identity of operators (56) and we can therefore deduce (57) from (58). In other words, every solution of the differential equation is an isobaric function of weight p.

More generally, whenever a substitution involving a disposable parameter leaves a function 'invariant', i.e. at most affected by a multiplier,

we may expect to be able to determine a partial differential equation satisfied by the function, by similarly differentiating the identity which states the invariance. This is certainly the case, for example, in the theory of the Invariants of Algebraic Forms.

As a simpler example consider the differential equation satisfied by a *function of differences*. We say that $f(x_1,..., x_n)$ is a function of differences of its arguments, if we can write identically

$$f(x_1,..., x_n) = \phi(x_1-x_2,..., x_r-x_s,...) \tag{59}$$

for some ϕ. More simply, such a function is defined by the identity

$$f(x_1+t,..., x_n+t) = f(x_1,..., x_n). \tag{60}$$

Evidently (60) is a consequence of (59) and we recover (59) at once by writing $t = -x_r$. This is legitimate, since t is a disposable parameter. Such a function of differences is thus invariant under the substitution

$$X_1 = x_1+t,..., \quad X_n = x_n+t. \tag{61}$$

The appropriate identity of differential operators is

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial X_1}+...+\frac{\partial}{\partial X_n}. \tag{62}$$

Applied to (60) this gives at once

$$\frac{\partial f}{\partial X_1}+...+\frac{\partial f}{\partial X_n} = 0.$$

Thus the differential equation characteristic of functions of differences is

$$\frac{\partial f}{\partial x_1}+...+\frac{\partial f}{\partial x_n} = 0. \tag{63}$$

The converse follows at once as in the preceding examples.

11. Interchange of dependent and independent variable

We have so far applied partial differentiation to prescribed functions of independent arguments. Dependent and independent variables have thus been assigned without liberty of choice. Consider now the case of n variables $x_1,..., x_n$ connected by a single relation

$$f(x_1,..., x_n) = 0. \tag{64}$$

In general, we may suppose that this equation can be solved† for any argument x_r in the form

$$x_r = x_r(x_1,..., x_{r-1}, x_{r+1},..., x_n). \tag{65}$$

† The limitations under which this is possible are discussed later in chapter XI.

If we further suppose that these new functions (65) are all differentiable, it becomes legitimate to speak of the $n(n-1)$ partial derivatives

$$\frac{\partial x_r}{\partial x_s} \qquad (r \neq s).$$

Similarly, if further differentiations are permissible, we may pass on to consider the $\frac{1}{2}n^2(n-1)$ derivatives of the second order

$$\frac{\partial^2 x_r}{\partial x_s \partial x_t},$$

where s, t may be the same but must differ from r, and so on.

These partial derivatives, as we should expect, are not all independent. Formulae may be obtained connecting them with one another and also with the partial derivatives of f in its n arguments.

If we apply the formula of the total differential to the defining equation (64), we have

$$0 = \frac{\partial f}{\partial x_1} Dx_1 + \dots + \frac{\partial f}{\partial x_n} Dx_n.$$

We form $\partial x_r/\partial x_s$ by keeping constant every argument except x_r, x_s, i.e. by equating to zero every Dx except Dx_r, Dx_s. This gives

$$0 = \frac{\partial f}{\partial x_r} Dx_r + \frac{\partial f}{\partial x_s} Dx_s,$$

and so
$$\frac{Dx_r}{Dx_s} = -\frac{\partial f}{\partial x_s}\Big/\frac{\partial f}{\partial x_r}.$$

In proper notation we therefore have

$$\frac{\partial x_r}{\partial x_s} = -\frac{\partial f}{\partial x_s}\Big/\frac{\partial f}{\partial x_r} \qquad (r \neq s). \tag{66}$$

In particular we deduce that

$$\left.\begin{aligned} \frac{\partial x_1}{\partial x_2}\frac{\partial x_2}{\partial x_1} &= 1 \\ \frac{\partial x_1}{\partial x_2}\frac{\partial x_2}{\partial x_3}\frac{\partial x_3}{\partial x_1} &= -1 \\ \text{and generally} \quad \frac{\partial x_1}{\partial x_2}\cdots\frac{\partial x_{r-1}}{\partial x_r}\frac{\partial x_r}{\partial x_1} &= (-1)^r \end{aligned}\right\}. \tag{67}$$

These results are in sharp contrast to those which we obtain, if there is only *one* degree of freedom in the field $(x_1,\dots,x_n)$, i.e. if the n variables

are connected not by one but by $n-1$ independent relations. The differentiations are then total instead of partial and we have the formulae

$$\left.\begin{aligned} \frac{dx_1}{dx_2}\frac{dx_2}{dx_1} &= 1 \\ \frac{dx_1}{dx_2}\frac{dx_2}{dx_3}\frac{dx_3}{dx_1} &= 1 \\ \text{and generally} \qquad \frac{dx_1}{dx_2}\cdots\frac{dx_{r-1}}{dx_r}\frac{dx_r}{dx_1} &= 1 \end{aligned}\right\}, \tag{68}$$

in which the alteration of sign on the right is significant.

To obtain the second derivative $\partial^2 x_r/\partial x_s \partial x_t$ we have to differentiate (66) partially in x_t on the supposition that, on the right, the dependent variable x_r has already been replaced in terms of $x_1, \ldots, x_{r-1}, x_{r+1}, \ldots, x_n$ by its equivalent expression (65). In other words, on the right of (66) we suppose x_t to enter not only explicitly but also through x_r. We therefore have

$$\frac{\partial^2 x_r}{\partial x_s \partial x_t} = \left\{\frac{\partial}{\partial x_t} + \frac{\partial x_r}{\partial x_t}\frac{\partial}{\partial x_r}\right\}\left(-\frac{\partial f}{\partial x_s}\Big/\frac{\partial f}{\partial x_r}\right),$$

where $\partial x_r/\partial x_t$ means the appropriate derivative of (65) and so is given by the formula corresponding to (66).

Hence, performing the differentiations and substituting for $\partial x_r/\partial x_t$, we have at length

$$\frac{\partial^2 x_r}{\partial x_s \partial x_t} = \frac{-f_{rr} f_s f_t + f_r (f_{rs} f_t + f_{rt} f_s) - f_r^2 f_{st}}{f_r^3}. \tag{69}$$

It is evidently immaterial to the argument whether x_s, x_t are the same or different. The formula for $\partial^2 x_r/\partial x_s^2$ is therefore sufficiently covered by (69).

In similar fashion we may obtain the typical formula for derivatives of the third order

$$\frac{\partial^3 x_r}{\partial x_s \partial x_t \partial x_u} = \{(f_r f_{rrr} - 3f_{rr}^2) f_s f_t f_u + 3 f_r f_{rr} \sum f_{rs} f_t f_u -$$

$$-f_r^2 \sum (f_{rrs} f_t + 2 f_{rs} f_{rt} + f_{rr} f_{st}) f_u + f_r^3 \sum (f_{rst} f_u + f_{rs} f_{tu}) - f_r^4 f_{stu}\}/f_r^5, \tag{70}$$

where the summations $\sum$ extend over the suffixes s, t, u. It is clear that these formulae soon become unmanageable.

Suitable algebraic elimination of the partial derivatives of f between these formulae will evidently give the various relations that connect the partial derivatives of x_r with those of another variable x_s. But in practice we most simply express the derivatives of x_r in terms of those of x_s by supposing the relation (64) to be written in the special form

$$x_s(x_1, \ldots, x_{s-1}, x_{s+1}, \ldots, x_n) - x_s = 0.$$

Then $f_s = -1$ and all the derivatives of f_s vanish, while the derivatives of f in the other arguments become the derivatives of x_s in those arguments. As an illustration consider a set of three variables x, y, z connected by a single relation, expressing the second derivatives of y in terms of the derivatives of z. From (69) we derive at once the three formulae

$$\left.\begin{aligned} \frac{\partial^2 y}{\partial x^2} &= -\left\{\frac{\partial^2 z}{\partial x^2}\left(\frac{\partial z}{\partial y}\right)^2 - 2\frac{\partial^2 z}{\partial x \partial y}\frac{\partial z}{\partial x}\frac{\partial z}{\partial y} + \frac{\partial^2 z}{\partial y^2}\left(\frac{\partial z}{\partial x}\right)^2\right\}\Big/\left(\frac{\partial z}{\partial y}\right)^3 \\ \frac{\partial^2 y}{\partial x \partial z} &= -\left\{\frac{\partial^2 z}{\partial x \partial y}\frac{\partial z}{\partial y} - \frac{\partial^2 z}{\partial y^2}\frac{\partial z}{\partial x}\right\}\Big/\left(\frac{\partial z}{\partial y}\right)^3 \\ \frac{\partial^2 y}{\partial z^2} &= -\frac{\partial^2 z}{\partial y^2}\Big/\left(\frac{\partial z}{\partial y}\right)^3 \end{aligned}\right\}. \tag{71}$$

12. Inadequacy of the notation

If we go farther and seek to consider the mutual partial derivatives of variables connected by *two* relations

$$f(x_1,\ldots,x_n) = 0, \qquad g(x_1,\ldots,x_n) = 0, \tag{72}$$

we are met by difficulties. Let us suppose that we wish to form the partial derivative of x_r with respect to x_s. We cannot, as we have hitherto done, retain x_r, x_s alone as variables and regard every other x as a constant For then x_r, x_s are defined in terms of constants by the two relations (72) and so are themselves constants. We must accordingly retain as a variable at least one other x, say x_t.

In the total differentials of (72)

$$f_1\,Dx_1 + \ldots + f_n\,Dx_n = 0,$$
$$g_1\,Dx_1 + \ldots + g_n\,Dx_n = 0$$

we now put every Dx zero except Dx_r, Dx_s, Dx_t and so get

$$f_r\,Dx_r + f_s\,Dx_s + f_t\,Dx_t = 0,$$
$$g_r\,Dx_r + g_s\,Dx_s + g_t\,Dx_t = 0.$$

Elimination of the irrelevant Dx_t gives

$$(f_r g_t - f_t g_r)\,Dx_r + (f_s g_t - f_t g_s)\,Dx_s = 0,$$

i.e.
$$\frac{\partial x_r}{\partial x_s} = -\frac{f_s g_t - f_t g_s}{f_r g_t - f_t g_r}.$$

The value of $\partial x_r/\partial x_s$ in these circumstances thus depends on our choice of subsidiary variable x_t. We could indicate this in the notation, if we wished, by introducing some such symbol as

$$\left(\frac{\partial x_r}{\partial x_s}\right)_t.$$

These difficulties evidently increase with derivatives of higher order, for we have a free choice of subsidiary variable at each differentiation. They increase too with increase in the number of relations connecting the n variables x_r. If there are m such relations, then to be able to form a derivative, say of x_r with respect to x_s, we must regard $n-m-1$ of the other variables as constants. But still more generally, we could reduce the degrees of freedom to one (to form a derivative), if we superimposed on the given m relations any $n-m-1$ other relations, or prescribed that any $n-m-1$ other functions should remain constant. At this stage the notion of 'partial' differentiation has become so diffuse as scarcely to be of value.

In practice, partial differentiation is most usefully retained for a function of independent arguments and may best be regarded, not particularly as determining rates of change, but as defining auxiliary functions intimately associated with the given function Ordinary or total differentiation, on the other hand, is precisely concerned with measuring relative rates of change when, in a particular problem, the available degrees of freedom have been reduced to one

Moreover, for a function f of n arguments, the associated functions, namely the partial derivatives $f_1, f_2, \ldots, f_n$, are best spoken of as the partial derivatives *in the first, second,..., nth argument*. We can then assign to these n arguments any special values we wish, in particular values not necessarily independent, and so speak without confusion of such partial derivatives as
$$f_1(x,x,x), \qquad f_3(y-z,z-x,x-y).$$

WORKED EXAMPLE

If for every (x,y) in some interval
$$\lim_{h,k\to 0}\frac{f(x+h,y+k)-f(x+h,y)-f(x,y+k)+f(x,y)}{hk}=0,$$
then, throughout the interval, $f(x,y)$ is of the form $\phi(x)+\psi(y)$.

We have $\quad |f(x\pm h,y\pm k)-f(x\pm h,y)-f(x,y\pm k)+f(x,y)| < hk\epsilon,$

if $$0 < h,k < \text{some } \eta(\epsilon,x,y).$$

Hence, if h', k' are also positive and less than this η, we shall have simultaneously
$$|f(x-h',y+k)-f(x-h',y)-f(x,y+k)+f(x,y)| < h'k\epsilon,$$
$$|f(x+h,y-k')-f(x+h,y)-f(x,y-k')+f(x,y)| < hk'\epsilon,$$
$$|f(x-h',y-k')-f(x-h',y)-f(x,y-k')+f(x,y)| < h'k'\epsilon.$$

By combining these four inequalities we have
$$|f(x+h,y+k)-f(x+h,y-k')-f(x-h',y+k)+f(x-h',y-k')| < (h+h')(k+k')\epsilon.$$

In other words, if (x_1, y_1), (x_2, y_2) are any two points within the square $(x \pm \eta, y \pm \eta)$, where $\eta = \eta(\epsilon, x, y)$, the inequality

$$|f(x_1,y_1)-f(x_1,y_2)-f(x_2,y_1)+f(x_2,y_2)| < |x_1-x_2|\cdot|y_1-y_2|\epsilon \tag{1}$$

is satisfied.

Fix x, y, ϵ and consider the set of numbers $\{\eta\}$ that secure this inequality. The set is bounded, because we have restricted ourselves to points of an interval. Thus the set will have a greatest number or an upper bound $\eta_0(\epsilon, x, y)$, which, for a fixed ϵ, defines $\eta_0(x,y)$ a function of (x,y). We proceed to show that it is a continuous function.

If now, h, k being of either sign, $|h|$, $|k| < \eta_0(x,y)$ and η' is the smaller of $\eta_0-|h|$, $\eta_0-|k|$, then the square

$$(x+h\pm\eta', y+k\pm\eta')$$

lies wholly within the square $(x\pm\eta_0, y\pm\eta_0)$ and hence the inequality (1) holds throughout the smaller square. Thus, by the definition of $\eta_0(x,y)$,

$$\eta' \leq \eta_0(x+h, y+k),$$

i.e.
$$\eta_0(x,y)-\eta_0(x+h,y+k) \leq \max(|h|,|k|).$$

Similarly, by interchanging (x,y) and $(x+h,y+k)$ we can prove that

$$\eta_0(x+h,y+k)-\eta_0(x,y) \leq \max(|h|,|k|).$$

Hence
$$|\eta_0(x+h,y+k)-\eta_0(x,y)| < \epsilon, \text{ if } |h|,|k| < \epsilon.$$

Accordingly, $\eta_0(x,y)$ is continuous in the interval and has therefore in the interval a least value, which, being its value at some point of the interval, is greater than zero. Call this least value $\delta(\epsilon)$, it is independent of x, y, but presumably depends on ϵ. Thus (1) is secured, if $|x_1-x_2|$, $|y_1-y_2| < \delta(\epsilon)$, or, in other notation,

$$|f(x+h,y+k)-f(x+h,y)-f(x,y+k)+f(x,y)| < |h|\,|k|\epsilon, \tag{2}$$

if
$$|h|,|k| < \delta(\epsilon).$$

Change the signs of h, k independently and add the four resulting inequalities.† We get that

$$|f(x+h,y+k)-f(x+h,y-k)-f(x-h,y+k)+f(x-h,y-k)| < |2h|\,|2k|\epsilon,$$

if
$$|h|,|k| < \delta(\epsilon)$$

or, if we write $x+h$, $y+k$ for x, y,

$$|f(x+2h,y+2k)-f(x+2h,y)-f(x,y+2k)+f(x,y)| < |2h|\,|2k|\epsilon,$$

if
$$|2h|,|2k| < 2\delta(\epsilon).$$

This, in effect, is (2) with 2δ replacing δ. A second application of the process enables us to replace 2δ by 4δ and so on. We may thus at length remove the restriction on $|h|$, $|k|$ and so have for *every* h, k

$$|f(x+h,y+k)-f(x+h,y)-f(x,y+k)+f(x,y)| < |h|\,|k|\epsilon.$$

Since h, k are now independent of ϵ, we may take the limit $\epsilon \to 0$ and get, for every (x,y), $(x+h,y+k)$ of the given interval,

$$f(x+h,y+k)-f(x+h,y)-f(x,y+k)+f(x,y)=0.$$

If (a,b) is a fixed point of the interval, write $h=a-x$, $k=b-y$ and we have throughout the interval

$$f(x,y)=f(a,y)+f(x,b)-f(a,b),$$

i.e. $f(x,y)$ is of the stated form $\phi(x)+\psi(y)$.

† As in the first section of this proof

EXAMPLES VI

1. Discuss, at the origin, the existence of the partial derivatives f_x, f_y and the continuity and differentiability of $f(x,y)$ itself, when $f(x,y)$ has the following forms:

(i) $\sqrt{(x^2+y^2)}$, (ii) $\dfrac{x^p y^q}{x^{2r}+y^{2s}}$ $(p,q,r,s>0)$,

(iii) $xy\sin\dfrac{1}{xy}$, (iv) $xy\sin\dfrac{1}{x^2+y^2}$,

(v) $(x^2+y^2)\sin\dfrac{1}{x^2+y^2}$, (vi) $x^2y^2\log(x^2+y^2)$,

(vii) $xy\log(x^2+y^2)$, (viii) $xyI(x)I(y)$,

where $I(x)$ denotes the greatest integer in x

[In every case we suppose at the origin that $f(0,0)=0$ and in (iii) we extend the definition by $f(x,0)=0=f(0,y)$.]

2. Show that the function

$$f(x,y)=\frac{xy^2}{x^2+y^4},\qquad f(0,0)=0$$

is differentiable at the origin along every direction $y=kx$, but is not continuous there.

3. Show that, if $f_x(x,y), f_y(x,y)$ are differentiable at a point, then $f(x,y)$ itself is differentiable at the point

4. Prove that, if $f_x(x,y), f_y(x,y)$ are differentiable at a point, then near that point

$$\begin{aligned}f(x+\delta_1x+\delta_2x,\,y+\delta_1y+\delta_2y)-f(x+\delta_1x,\,y+\delta_1y)-f(x+\delta_2x,\,y+\delta_2y)+f(x,y)\\ =A\delta_1x\delta_2x+B\delta_1x\delta_2y+C\delta_1y\delta_2x+D\delta_1y\delta_2y,\end{aligned}$$

where A, B, C, D are functions of $x, y, \delta_1x, \delta_1y, \delta_2x, \delta_2y$ which converge as $\delta_1x, \delta_1y, \delta_2x, \delta_2y\to 0$.

Prove, conversely, that, if near a point the double difference is expressible in this form, then f_x, f_y exist near the point and are both differentiable there.

5. (i) If x, y, z be three angles connected by the relation

$$\frac{1-\cos^2x-\cos^2y-\cos^2z+2\cos x\cos y\cos z}{\sin^2x\sin^2y\sin^2z}=\text{constant},$$

prove that

$$\frac{\sin y\sin z\,dx}{\cos x-\cos y\cos z}+\frac{\sin z\sin x\,dy}{\cos y-\cos z\cos x}+\frac{\sin x\sin y\,dz}{\cos z-\cos x\cos y}=0$$

(ii) If the sides and angles of a spherical triangle vary subject to the condition

$$\frac{\sin A}{\sin a}=\frac{\sin B}{\sin b}=\frac{\sin C}{\sin c}=\text{constant},$$

prove that

$$\sec A\,da+\sec B\,db+\sec C\,dc=0,$$
$$\sec a\,dA+\sec b\,dB+\sec c\,dC=0$$

6. If $$\sqrt{z}=\frac{\sqrt{\{x(1+ay)(1+by)\}}+\sqrt{\{y(1+ax)(1+bx)\}}}{1-abxy},$$

and if x, y, z vary, while a, b remain constant, prove that

$$\frac{dx}{\sqrt{\{x(1+ax)(1+bx)\}}}+\frac{dy}{\sqrt{\{y(1+ay)(1+by)\}}}-\frac{dz}{\sqrt{\{z(1+az)(1+bz)\}}}=0.$$

7. Discuss the existence and equivalence at the origin of f_{xy}, f_{yx} for the following functions:

(i) $\dfrac{a_0x^4+4a_1x^3y+6a_2x^2y^2+4a_3xy^3+a_4y^4}{x^2+y^2}$,

(ii) $\dfrac{(ax^2+2bxy+cy^2)^3}{x^2+y^2}$, (iii) $\dfrac{(ax^2+2bxy+cy^2)^{\frac{3}{2}}}{(x^2+y^2)^{\frac{1}{2}}}$,

(iv) $\dfrac{x^4+(py+p'y^2+\dots)x^3+(qy^2+q'y^3+\dots)x^2+(ry^4+r'y^5+\dots)x+(sy^6+s'y^7+\dots)}{x^2+y^4}$,

where in each case we define $f(0,0)$ as the limit of $f(x,y)$ at the origin.

Discuss also the satisfaction of the conditions (27), (29) of § 7.

8. If
$$f(x,y)=\frac{x^5y-x^3y^2}{x^4+x^2y+y^2},\qquad f(0,0)=0,$$
show that at the origin $f_{xy}\neq f_{yx}$, but that f_x is differentiable and f_y continuous at the origin. Discuss the applicability of the conditions (29) of § 7.

9. By considering the function
$$f(x,y)=\frac{x^2y^2}{x^2+y^2},\qquad f(0,0)=0,$$
show, in the notation of § 7, that the equivalence of f_{xy}, f_{yx} is insufficient to secure the continuity of
$$\phi(h,k)\quad\text{at}\quad(0,0)$$

By considering a function of the form
$$f(x,y)=yF(x),$$
show that the continuity of $\phi(h,k)$ at $(0,0)$ is insufficient to secure the continuity of f_{xy} at (ξ,η).

10. Show that
$$\phi(h,k,x,y)=\frac{f(x+h,y+k)-f(x+h,y)-f(x,y+k)+f(x,y)}{hk}$$
is continuous in (h,k,x,y) at $(0,0,\xi,\eta)$, if *either* of f_{xy}, f_{yx} exists and is continuous at (ξ,η); and conversely that, if ϕ is continuous at the point stated, then *both* f_{xy}, f_{yx} exist, if f_x, f_y exist at (ξ,η) and are continuous at (ξ,η), if f_{xy}, f_{yx} exist near (ξ,η).

11. Show that $\theta(a,h)$ defined by Taylor's remainder-formula
$$f(a+h)=\sum_{r=0}^{n-1}\frac{h^r}{r!}f_r(a)+\frac{h^n}{n!}f_n(a+\theta h)$$
satisfies the partial differential equation
$$\frac{h^2\left(\dfrac{\partial\theta}{\partial a}-\dfrac{\partial\theta}{\partial h}\right)+h(1-\theta)}{nh\dfrac{\partial\theta}{\partial a}+n}=\frac{f(a+h)-\displaystyle\sum_{r=0}^{n}\frac{h^r}{r!}f_r(a)}{f'(a+h)-\displaystyle\sum_{r=0}^{n-1}\frac{h^r}{r!}f_{r+1}(a)},$$
supposing that $f_{n+1}(x)$ exists at and near $x=a$.

12. Show that the function $\theta(a,b,h,k)$ defined by the generalized formula of the mean
$$f(a+h,b+k)=f(a,b)+hf_1(a+\theta h,b+\theta k)+kf_2(a+\theta h,b+\theta k)$$

satisfies the partial differential equation

$$\frac{h\dfrac{\partial\theta}{\partial h}+k\dfrac{\partial\theta}{\partial k}+\theta}{h\dfrac{\partial\theta}{\partial a}+k\dfrac{\partial\theta}{\partial b}+1}=\frac{hf_1(a+h,b+k)+kf_2(a+h,b+k)-f(a+h,b+k)+f(a,b)}{hf_1(a+h,b+k)+kf_2(a+h,b+k)-hf_1(a,b)-kf_2(a,b)},$$

where the suffixes indicate differentiation of f in its first and second argument respectively, and where $f_1(x,y)$, $f_2(x,y)$ are supposed differentiable in the interval $(a,a+h;b,b+k)$.

13. Show that, granted differentiability, the function $z(x,y)$ defined by the formula of the mean

$$f(x)-f(y)=(x-y)f'(z)$$

satisfies the partial differential equation

$$(x-y)^2\{z_x z_y(z_y z_{xxy}-z_x z_{xyy})-z_{xy}(z_{xx}z_y^2-z_{yy}z_x^2)\}-$$
$$-(x-y)\{z_{xx}z_y^3-z_x z_y z_{xy}(z_x+z_y)+z_x^3 z_{yy}\}+z_x z_y(z_x^2-z_y^2)=0.$$

Obtain the general solution of this differential equation.

14. If $f(x,y)$ be a polynomial, show that the necessary and sufficient condition that it be symmetrical in x,y is that

$$\sum_{r=1}^{\infty}\frac{(y-x)^{r-1}}{r!}\left(\frac{\partial}{\partial x}-\frac{\partial}{\partial y}\right)^r f=0.$$

15. If f be a function of two arguments such that, at (a,b), f_1, f_2 are differentiable and $f_{11}f_{22}>f_{12}^2$, prove that, as $h,k\to 0$,

$$\frac{f(a+h,b+k)-f(a,b)-hf_1(a,b)-kf_2(a,b)}{h^2f_{11}+2hkf_{12}+k^2f_{22}}\to\tfrac{1}{2}.$$

16. The number θ $(0<\theta<1)$ is defined by the Taylor's formula

$$f(a+h,b+k)=f(a,b)+hf_1(a+\theta h,b+\theta k)+kf_2(a+\theta h,b+\theta k).$$

Prove that, if f_1, f_2 are differentiable at (a,b) and if $h,k\to 0$ so that

$$(h^2+k^2)/\{h^2f_{11}(a,b)+2hkf_{12}(a,b)+k^2f_{22}(a,b)\}$$

is bounded, then $\theta\to\frac{1}{2}$.

If f_{11}, f_{12}, f_{22} are differentiable at (a,b) and if $h,k\to 0$ so that h/k converges to a root of the equation

$$t^2f_{11}(a,b)+2tf_{12}(a,b)+f_{22}(a,b)=0$$

which is not also a root of the equation

$$t^3f_{111}(a,b)+3t^2f_{112}(a,b)+3tf_{122}(a,b)+f_{222}(a,b)=0,$$

while
$$\frac{h^2f_{11}(a,b)+2hkf_{12}(a,b)+k^2f_{22}(a,b)}{h^3f_{111}(a,b)+3h^2kf_{112}(a,b)+3hk^2f_{122}(a,b)+k^3f_{222}(a,b)}\to c,$$

prove that
$$\theta\to+\sqrt{(c^2+c+\tfrac{1}{3})}-c,\quad\text{if}\quad c>-\tfrac{1}{3},$$
$$\to-\sqrt{(c^2+c+\tfrac{1}{3})}-c,\quad\text{if}\quad c<-\tfrac{2}{3}.$$

17. If $f(x,y,z)$ be a homogeneous function, not necessarily algebraic, of degree n, show that

(i) $f_x(x,y,z)$ is homogeneous of degree $n-1$;

(ii) if the equation $f(x,y,z)=0$ defines an implicit function $z(x,y)$, this function is of degree unity;

(iii) if $f(x, y, z)$ can be expressed as a function of f_x, f_y, f_z, it is of degree $n/(n-1)$ in these new arguments;

(iv) if x, y, z are themselves homogeneous functions of u, v, w of degree p, then f is homogeneous of degree np in u, v, w.

18. If $f(x, y)$ is homogeneous of degree n, show that

$$\text{(i)} \quad \sum_{r=0}^{m} (m\,!\,r) x^{m-r} y^{r} \frac{\partial^m f}{\partial x^{m-r} \partial y^{r}} = n(n-1)\dots(n-m+1)f,$$

$$\text{(ii)} \quad \sum_{r=0}^{m} (m\,!\,r) \frac{\partial^m (x^{m-r} y^r f)}{\partial x^{m-r} \partial y^{r}} = (n+2)(n+3)\dots(n+m+1)f,$$

the symbols $(m\,!\,r)$ denoting binomial coefficients.

19. If $f(x, y, z)$ is homogeneous of degree n, show that

$$\text{(i)} \quad \begin{vmatrix} f_{xx} & f_{xy} & f_{xz} & f_x \\ f_{xy} & f_{yy} & f_{yz} & f_y \\ f_{xz} & f_{yz} & f_{zz} & f_z \\ f_x & f_y & f_z & 0 \end{vmatrix} \div \begin{vmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{xy} & f_{yy} & f_{yz} \\ f_{xz} & f_{yz} & f_{zz} \end{vmatrix} = -\frac{n}{n-1} f,$$

$$\text{(ii)} \quad \frac{(n-1)(f_y^2 f_{zz} - 2 f_y f_z f_{yz} + f_z^2 f_{yy}) - nf(f_{yy} f_{zz} - f_{yz}^2)}{x^2}$$

is symmetrical in x, y, z.

20. Show that the algebraic function $z(x_1, \dots, x_n)$ defined by the implicit relation

$$\sum_{r=1}^{n} \frac{x_r^2}{z+a_r} = 1,$$

where the a's are constants, satisfies the differential equations

$$\sum_{r=1}^{n} \left(\frac{\partial z}{\partial x_r}\right)^2 = 2 \sum_{r=1}^{n} x_r \frac{\partial z}{\partial x_r}, \qquad \sum_{r=1}^{n} \frac{\partial^2 z}{\partial x_r^2} = \sum_{r=1}^{n} \frac{1}{x_r} \frac{\partial z}{\partial x_r}.$$

21. If the n variables $x_1, x_2, \dots, x_n$ are connected by a single relation, show that

$$\frac{\partial^2 x_2}{\partial x_1^2} \frac{\partial^2 x_3}{\partial x_2^2} \cdots \frac{\partial^2 x_1}{\partial x_n^2} = \frac{\partial^2 x_1}{\partial x_2^2} \frac{\partial^2 x_2}{\partial x_3^2} \cdots \frac{\partial^2 x_n}{\partial x_1^2}.$$

22. If z, u be functions of x, y such that the equation

$$\sum_{r=0}^{p} (p\,!\,r) u^{p-r} \frac{\partial^p z}{\partial x^{p-r} \partial y^{r}} = 0$$

is true for $p = n-1, n$, then it is true for $p = n+1, n+2, \dots$.

23. If x, y, z are connected by a single relation, show that

$$\frac{\partial}{\partial y} \log\left(\frac{\partial z}{\partial x} \Big/ \frac{\partial z}{\partial y}\right) = \frac{\partial}{\partial y} \log\left(\frac{\partial x}{\partial z} \Big/ \frac{\partial x}{\partial y}\right),$$

where $\partial/\partial y$ is properly interpreted on the two sides of the equation.

24. If x, y, z are connected by a single relation, prove that any determinant of the set

$$(z_x)^{-\frac{3}{2}} \left\| \begin{matrix} z_x z_{xx} & z_{xx} & z_y z_{xx} + (z_x - 1) z_{xy} & z_y z_{xy} \\ z_x^2 + 1 & 2 z_x & 2 z_y (z_x - 1) & z_y^2 \end{matrix} \right\|$$

is unaltered by interchange of x and z.

25. If $f(x,y,z)=0$, show that the minors of f_{xx}, f_{xy}, f_{yy} in

$$\begin{vmatrix} f_{xx} & f_{xy} & f_{xz} & f_x \\ f_{xy} & f_{yy} & f_{yz} & f_y \\ f_{xz} & f_{yz} & f_{zz} & f_z \\ f_x & f_y & f_z & 0 \end{vmatrix}$$

are respectively $\qquad z_{yy}f_z^3, \quad -z_{xy}f_z^3, \quad z_{xx}f_z^3.$

Show that

$$\frac{z_{xx}z_{yy}-z_{xy}^2}{(z_x z_y)^{\frac{4}{3}}}$$

is symmetrical in x, y, z.

26. If x, y are connected by the implicit relation $f(x,y)=0$, show that

$$\left(\frac{\partial f}{\partial x}\right)^3\frac{\partial^3 f}{\partial y^3}-3\left(\frac{\partial f}{\partial x}\right)^2\frac{\partial f}{\partial y}\frac{\partial^3 f}{\partial x\partial y^2}+3\frac{\partial f}{\partial x}\left(\frac{\partial f}{\partial y}\right)^2\frac{\partial^3 f}{\partial x^2\partial y}-\left(\frac{\partial f}{\partial y}\right)^3\frac{\partial^3 f}{\partial x^3}$$
$$=\left(\frac{\partial f}{\partial y}\right)^4\left\{\frac{d^3y}{dx^3}+3\frac{d^2y}{dx^2}\frac{d}{dx}\left(\log\frac{\partial f}{\partial y}\right)\right\}$$
$$=\frac{\left(\frac{\partial f}{\partial y}\right)^6\frac{d}{dx}\left(\frac{\partial f}{\partial x}\right)\frac{d^3y}{dx^3}-\left(\frac{\partial f}{\partial x}\right)^6\frac{d}{dy}\left(\frac{\partial f}{\partial y}\right)\frac{d^3x}{dy^3}}{\left(\frac{\partial f}{\partial y}\right)^2\frac{\partial^2 f}{\partial x^2}-\left(\frac{\partial f}{\partial x}\right)^2\frac{\partial^2 f}{\partial y^2}}.$$

27. If x, y, z are connected by a single relation, show that the expressions

$$\left(\frac{\partial x}{\partial y}\frac{\partial^3 x}{\partial z^3}-3\frac{\partial^2 x}{\partial y\partial z}\frac{\partial^2 x}{\partial z^2}\right)\Big/\left\{\left(\frac{\partial x}{\partial z}\right)^2+\left(\frac{\partial x}{\partial z}\right)^3\right\},$$
$$\left\{\frac{\partial^3 x}{\partial z^3}\left(\frac{\partial x}{\partial y}\right)^2-3\frac{\partial^2 x}{\partial y\partial z}\frac{\partial^2 x}{\partial z^2}\frac{\partial x}{\partial y}-2\frac{\partial x}{\partial z}\left(\frac{\partial^3 x}{\partial y\partial z^2}\frac{\partial x}{\partial y}-3\frac{\partial^2 x}{\partial y^2}\frac{\partial^2 x}{\partial z^2}\right)\right\}\Big/\left(\frac{\partial x}{\partial z}\right)^3$$

are unaltered by interchange of x and z.

28. If x, y, z, u are connected by *two* relations, and if the notation

$$\left(\frac{\partial y}{\partial x}\right)_z$$

means that z is kept constant in forming the partial derivative, show that

$$\left(\frac{\partial y}{\partial x}\right)_u\left(\frac{\partial x}{\partial y}\right)_z\left(\frac{\partial u}{\partial z}\right)_x\left(\frac{\partial z}{\partial u}\right)_y=1,$$
$$\left(\frac{\partial y}{\partial x}\right)_u\left(\frac{\partial x}{\partial y}\right)_z+\left(\frac{\partial z}{\partial x}\right)_u\left(\frac{\partial x}{\partial z}\right)_y=1,$$
$$\left(\frac{\partial y}{\partial u}\right)_x\left(\frac{\partial x}{\partial z}\right)_u=\left(\frac{\partial y}{\partial u}\right)_z\left(\frac{\partial x}{\partial z}\right)_y=-\left(\frac{\partial y}{\partial z}\right)_x\left(\frac{\partial x}{\partial u}\right)_z=-\left(\frac{\partial y}{\partial z}\right)_u\left(\frac{\partial x}{\partial u}\right)_y.$$

29. If $f(x,y)$, $g(x,y)$ are each differentiable in (x,y), show that, as $u, v\to 0$,

$$\begin{vmatrix} f(x+u,y+v) & g(x+u,y+v) & 1 \\ f(x-v,y+u) & g(x-v,y+u) & 1 \\ f(x,y) & g(x,y) & 1 \end{vmatrix}\div(u^2+v^2)\to\frac{\partial(f,g)}{\partial(x,y)}.$$

VII

INDETERMINATE FORMS. APPROXIMATIONS

1. Indeterminate forms

IT frequently happens in Analysis that the definition of a function breaks down at an isolated point, although the function is still convergent at the point. In such a case we may conveniently annex the point to the domain of the function by assigning the limit at the point as the value there, and the function then becomes continuous at the point. But, even if we do not wish to enlarge the domain in this way, it is generally desirable to establish the fact of convergence and to evaluate the limit.

Such a state of affairs arises most simply in the case of a function defined in the form $f(x)/g(x)$, where $f(x)$, $g(x)$ have a common (but isolated) zero at some point ξ of their common domain, which we may suppose to be a complete interval. Direct substitution, $x = \xi$, then gives the meaningless form $0/0$, although the function is defined everywhere near $x = \xi$. The form or value of such a function at such a point is said to be *indeterminate*, and the discussion of the limit

$$\lim_{x\to\xi} f(x)/g(x)$$

becomes of interest.

The most elementary example of such an indeterminate form is given by the ratio of two polynomials such as

$$(x^2-\xi^2)/(x-\xi). \tag{1}$$

If $x \neq \xi$, we may cancel the factor $x-\xi$ and reduce the function to the form

$$x+\xi \tag{2}$$

Thus *everywhere save at* $x = \xi$ the functions (1), (2) are identical. At $x = \xi$ itself the first function is undefined, but the second function is still defined and is continuous. As we pass to the limit $x \to \xi$, the identity of the two functions persists, since the point $x = \xi$ is expressly excluded from consideration in forming the limit. Thus the continuity of (2) at $x = \xi$ secures the convergence of (1) at that point. Precisely, we have

$$\lim_{x\to\xi}\frac{x^2-\xi^2}{x-\xi} = \lim_{x\to\xi}(x+\xi) = 2\xi. \tag{3}$$

This algebraic reduction of an indeterminate form by cancellation of common factors does not take us very far with transcendental functions

and we must supplement it by other methods. If the functions are differentiable, the Differential Calculus contributes a very convenient procedure, which we now consider.

2. L'Hospital's rule

Since by hypothesis $f(\xi) = 0 = g(\xi)$, we may write

$$f(x)/g(x) = \frac{f(x)-f(\xi)}{x-\xi}\Big/\frac{g(x)-g(\xi)}{x-\xi}.$$

Passage to the limit $x \to \xi$, on the right, gives

$$\lim_{x\to\xi} f(x)/g(x) = f'(\xi)/g'(\xi). \tag{4}$$

This is the characteristic result bearing the name of l'Hospital. It presupposes not only the existence of $f'(\xi)$, $g'(\xi)$ but also the non-vanishing of $g'(\xi)$.

If $g'(\xi) = 0$ but $f'(\xi) \neq 0$, we invert the fractions and obtain the corresponding limit

$$\lim_{x\to\xi} g(x)/f(x) = g'(\xi)/f'(\xi). \tag{5}$$

In this case it makes for conciseness to retain the limit (4) writing conventionally $$\lim_{x\to\xi} f(x)/g(x) = f'(\xi)/g'(\xi) = \infty.$$

If $f'(\xi)$, $g'(\xi)$ both vanish, the formula (4) fails, for it replaces one indeterminate form $0/0$ by another. Let us suppose more generally that possibly further derivatives of $f(x)$, $g(x)$ vanish at ξ. We then employ the following generalization of (4):

(6) *If $f'(\xi), g'(\xi),\ldots, f_{n-1}(\xi), g_{n-1}(\xi)$ exist and vanish and $f_n(\xi)$, $g_n(\xi)$ exist but do not both vanish, then*

$$\lim_{x\to\xi} f(x)/g(x) = f_n(\xi)/g_n(\xi)$$

For, since the derivatives $f'(\xi), g'(\xi),\ldots, f_{n-1}(\xi), g_{n-1}(\xi)$ all vanish, we can write

$$\frac{f(x)}{g(x)} = \frac{f(x)-\sum_{r=0}^{n-1}(x-\xi)^r f_r(\xi)/r!}{g(x)-\sum_{r=0}^{n-1}(x-\xi)^r g_r(\xi)/r!}.$$

As $x \to \xi$, the limit on the right, by chapter V (33), is $f_n(\xi)/g_n(\xi)$, if $g_n(\xi) \neq 0$. If $g_n(\xi) = 0$, $f_n(\xi) \neq 0$, we may write conventionally, as before, $f_n(\xi)/g_n(\xi) = \infty$. Hence under the conditions of (6) we have the required result $$\lim_{x\to\xi} f(x)/g(x) = f_n(\xi)/g_n(\xi).$$

In practice, we state the result of (6) in the form of a rule: to deter-

mine the limit of an expression $f(x)/g(x)$ at a point at which it assumes the indeterminate form 0/0, replace both numerator and denominator by their derivatives; if the new fraction $f'(x)/g'(x)$ still has the form 0/0 at the point, again apply the rule, replacing the new fraction by $f''(x)/g''(x)$, and so on.

The rule fails, if, up to the limits of differentiability, all the derivatives vanish at the point. Thus, if we write

$$\left.\begin{aligned} f(x) &= x^{\frac{3}{2}} \sin(\log x^2) \\ g(x) &= x^{\frac{3}{2}} \end{aligned}\right\}, \tag{7}$$

then $f'(0) = 0 = g'(0)$, but $f''(0)$, $g''(0)$ do not exist. Actually

$$f(x)/g(x) = \sin(\log x^2),$$

which does not converge at $x = 0$.

But, if we write

$$\left.\begin{aligned} f(x) &= x^2 \sin(\log x^2) \\ g(x) &= x^{\frac{3}{2}} \end{aligned}\right\}, \tag{8}$$

then again $f'(0) = 0 = g'(0)$, and $f''(0)$, $g''(0)$ do not exist. In this case, however, $f(x)/g(x) = x^{\frac{1}{2}} \sin(\log x^2) \to 0$ at $x = 0$.

The rule similarly fails, if $f(x)$, $g(x)$ are differentiable to any order but every derivative vanishes at $x = \xi$. Such a pair of functions is

$$\left.\begin{aligned} f(x) &= \exp\{-(x-\xi)^{-2}\} \\ g(x) &= \exp\{-\operatorname{cosec}^2(x-\xi)\} \end{aligned}\right\}. \tag{9}$$

Actually the limit is $\sqrt[3]{e}$.

In practice, these cases of failure are exceptional and for most purposes the rule embodied in (6) is exact and sufficient.

3. Alternative form of the rule

In actual working, however, the repeated differentiations easily become laborious and it is more convenient to be able to intersperse differential with algebraic methods. For this purpose we restate and elaborate the rule in the following form:

(10) *If $f(\xi)$, $g(\xi)$ vanish† and if $f'(x)$, $g'(x)$ do not vanish simultaneously in every neighbourhood of ξ, then*

$$\lim_{x\to\xi} f(x)/g(x) = \lim_{x\to\xi} f'(x)/g'(x),$$

if the second limit exists. More generally,

$$H > f(x)/g(x) > K \quad \textit{near } \xi,$$

if $$H > f'(x)/g'(x) > K \quad \textit{near } \xi.$$

† It is, of course, sufficient, if $f(x)$, $g(x) \to 0$ as $x \to \xi$. We are not, in point of fact, really concerned with values *at* ξ.

For, by the fractional formula of the mean, chapter IV (48),

$$\frac{f(x)}{g(x)} = \frac{f(x)-f(\xi)}{g(x)-g(\xi)} = \frac{f'(x_1)}{g'(x_1)},$$

where $x \gtrless x_1 \gtrless \xi$, unless $f'(x), g'(x)$ vanish simultaneously within the interval (x, ξ). Now if there is a neighbourhood of ξ in which $f'(x)$, $g'(x)$ have no common zero, we may proceed to the limit $x \to \xi$ through values of x lying entirely within this neighbourhood. Hence also $x_1 \to \xi$ through values confined to this neighbourhood. The conditions of chapter IV (48) are satisfied and we have

$$\lim_{x\to\xi} f(x)/g(x) = \lim_{x\to\xi} f'(x)/g'(x),$$

if the latter limit exists. The second part of (10) follows in an exactly similar fashion.

As an example of the exceptional case write

$$\left.\begin{aligned} f(x) &\equiv x^2\sin(2\log x)+x^2\cos(2\log x) \\ g(x) &\equiv x\sin(\log x) \end{aligned}\right\} \quad (x > 0). \tag{11}$$

Then
$$f'(x) = 4x\cos(2\log x),$$
$$g'(x) = \cos(\log x)+\sin(\log x)$$

Hence, cancelling the common factor $\cos(\log x)+\sin(\log x)$, we have

$$f'(x)/g'(x) = 4x\{\cos(\log x)-\sin(\log x)\},$$

which converges to zero at the origin. Now

$$f(x)/g(x) = 2x\{\cos(\log x)-\sin(\log x)\}+x\operatorname{cosec}(\log x).$$

At the origin, in the expression on the right, $x\operatorname{cosec}(\log x)$ oscillates infinitely while the remainder of the expression converges to zero. Hence $f(x)/g(x)$ oscillates infinitely at the origin, although $f'(x)/g'(x)$ converges there. This discrepancy is, of course, accounted for by the cancellation of the factor $\cos(\log x)+\sin(\log x)$, which vanishes in every neighbourhood of the origin.

It is hardly necessary to point out that we cannot use (10) in the reverse direction to determine the limit of $f'(x)/g'(x)$ from that of $f(x)/g(x)$. For the theorem of the mean on which we rely enables us to equate $f(x)/g(x)$ and $f'(x_1)/g'(x_1)$ for every x *but not for every* x_1. This is confirmed by an example:

$$\left.\begin{aligned} f(x) &\equiv x^2\sin(x^{-1}) \\ g(x) &\equiv x \end{aligned}\right\}. \tag{12}$$

Here $$\lim_{x\to 0} f(x)/g(x) = \lim_{x\to 0} x\sin(x^{-1}) = 0,$$

but $$f'(x)/g'(x) = 2x\sin(x^{-1})-\cos(x^{-1}),$$

which oscillates finitely at the origin. The condition of (10), that $f'(x)$, $g'(x)$ have no common zeros infinitely near the origin, is satisfied, since $g'(x) = 1$.

The formula of the mean does, however, enable us to assert that, if the condition about common zeros is satisfied, $f'(x)/g'(x)$ does converge to the limit of $f(x)/g(x)$ *along some sequence of approach*. For, if we let x tend to ξ along some arbitrary sequence, the sequence of corresponding values of x_1 is such a sequence of approach. Thus, in (12), $f'(x)/g'(x)$ converges to the limit of $f(x)/g(x)$, namely zero, along the sequence of approach $$x_n = 2/(2n+1)\pi.$$

4. Comparison of the two rules. Indeterminacy at infinity

It is worth while to compare the two methods of evaluating a limit given in §§ 2, 3, respectively. In using the method of § 2 we are free from the vexatious condition regarding a succession of simultaneous zeros, but we are possibly committed to repeated differentiations. With the method of § 3 we are not concerned with the value of $f'(x)/g'(x)$ actually at $x = \xi$. It may itself be indeterminate, or $f'(\xi)$, $g'(\xi)$ may not even exist. It is sufficient that we should be able to evaluate the limit of $f'(x)/g'(x)$ at $x = \xi$ by any available method. Of course, if the conditions of (10) apply to the new form $f'(x)/g'(x)$, we may evidently set out to evaluate by a second differentiation and so on until we reach a determinate form $f_n(\xi)/g_n(\xi)$. So applied, the method of § 3 has no advantage over that of § 2. Its utility lies in the fact that the limit of $f'(x)/g'(x)$ may be obtained *by any available method*. We can thus employ algebraic devices at any stage, cancelling a factor common to numerator and denominator or, more generally, evaluating the limit as a product of limits.

Now, if a function $\phi(x)$ vanishes in every neighbourhood of ξ, so too does $\phi'(x)$, if it exists, for, by Rolle's theorem, it vanishes between every two zeros of $\phi(x)$. Hence, by induction, a similar result holds for every existing derived function of $\phi(x)$. Again, if $\phi(x)$ is convergent at ξ, it must vanish there, since it vanishes in every neighbourhood of ξ. Since, then, $\phi(x)-\phi(\xi)$ vanishes in every neighbourhood of ξ, we see that $\phi'(\xi)$ and hence, by induction, every derivative $\phi_n(\xi)$, is zero at ξ.

Thus, if $f'(x)$, $g'(x)$ vanish simultaneously in every neighbourhood of ξ, then also every existing $f_n(x)$, $g_n(x)$ vanishes, though not necessarily simultaneously, in every neighbourhood of ξ, and any existing $f_r(\xi)/g_r(\xi)$

must necessarily be of the form $0/0$. The only possibility of securing a determinate limit for $f_n(x)/g_n(x)$ is accordingly by the cancellation of a common factor that vanishes in every neighbourhood of ξ. If we forbid such a cancellation, we are safely within the provisions of (10). Hence

(13) *In evaluating an indeterminate form $0/0$ we may intermix the differential method of differentiating simultaneously numerator and denominator with the algebraic method of cancelling common factors, provided only that such factors do not vanish in every neighbourhood of the point of indeterminacy.*

Some authorities would take exception to extending the term 'convergent' to a function which was indeterminate at points infinitely near the point of convergence. The narrower interpretation, of course, removes all difficulties of sequences of simultaneous zeros. I have, however, preferred to adopt the wider view, as set out in chapter II § 1, that $f(x) \to A$ at ξ, if $|f(x)-A| < \epsilon$ at all points *within the domain of definition* at which $|x-\xi| <$ some $\delta(\xi, \epsilon)$. Moreover, in such an example as (11), the points of indeterminacy (other than ξ) of $f'(x)/g'(x)$ are isolated points and $f'(x)/g'(x)$ is convergent there. Accordingly, in the spirit of the opening paragraph of this chapter we could reasonably extend the domain of definition of $f'(x)/g'(x)$ to include these points.

We can extend the method of § 3 to reduce a similar indeterminacy at infinity.

Suppose that $$\lim_{x\to\infty} f(x) = 0 = \lim_{x\to\infty} g(x).$$

Write $$x \equiv z^{-1},$$

then $$\lim_{x\to\infty} f(x)/g(x) = \lim_{z\to 0} f(z^{-1})/g(z^{-1}),$$

if the second limit exists.

Under the restrictions of (10) or (13) we may replace the latter limit by $$\lim_{z\to 0} z^{-2}f'(z^{-1})/z^{-2}g'(z^{-1}) = \lim_{z\to 0} f'(z^{-1})/g'(z^{-1})$$
$$= \lim_{x\to\infty} f'(x)/g'(x),$$

if this last limit exists.

Thus we have, in analogy with (10),

(14) *If* $$\lim_{x\to\infty} f(x) = 0 = \lim_{x\to\infty} g(x),$$

then $$\lim_{x\to\infty} f(x)/g(x) = \lim_{x\to\infty} f'(x)/g'(x),$$

if the second limit exists, unless $f'(x)$, $g'(x)$ have an unbounded set of common zeros.

5. The indeterminate form ∞/∞

We now consider the case in which

$$\lim_{x\to\xi} f(x) = \infty = \lim_{x\to\xi} g(x).$$

Then, at $x = \xi$, $\qquad \{f(x)\}^{-1}/\{g(x)\}^{-1}$

has the indeterminate form $0/0$ and hence under the usual conditions

$$\lim_{x\to\xi} \{f(x)\}^{-1}/\{g(x)\}^{-1} = \lim_{x\to\xi} f'(x)\{f(x)\}^{-2}/g'(x)\{g(x)\}^{-2},$$

if the second limit exists.

This tells us that, if *both* the limits

$$\lim_{x\to\xi} f(x)/g(x), \qquad \lim_{x\to\xi} f'(x)/g'(x)$$

exist, they are equal, and suggests the possibility of applying to the form ∞/∞ the results obtained for the form $0/0$.

We must observe that by chapter IV (64), if $f(x)$ have an infinity at ξ, so also has $f'(x)$. Thus the analogy of the first rule (6) is inapplicable, since every $f_n(\xi)/g_n(\xi)$ is also indeterminate of the form ∞/∞. Moreover, even if the second rule (10) is applicable, the differentiations do nothing to remove the infinities. Their purpose is to simplify the functions in the hope that it may be possible to take out the infinity by algebraic methods.

The fundamental proposition for the form ∞/∞ is the following:

(15) *If* $$\lim_{x\to\xi} f(x) = \infty = \lim_{x\to\xi} g(x)$$

and if $f'(x)$, $g'(x)$ do not vanish simultaneously in every neighbourhood of ξ, then

$$\lim_{x\to\xi} f(x)/g(x) = \lim_{x\to\xi} f'(x)/g'(x)$$

if the latter limit exists. More generally

$$H > f(x)/g(x) > K \quad \textit{near } \xi,$$

if $$H > f'(x)/g'(x) > K \quad \textit{near } \xi.$$

The enunciation presupposes that $f'(x)$, $g'(x)$ both exist throughout some open neighbourhood of ξ, so that ξ is an isolated infinity of $f(x), g(x)$.

Suppose now that $f'(x)/g'(x) \to A$ as $x \to \xi$, so that

$$|f'(x)/g'(x) - A| < \tfrac{1}{3}\epsilon, \quad \text{if } |x-\xi| < \text{some } \eta(\epsilon). \tag{16}$$

By hypothesis there is a neighbourhood of ξ, say $|x-\xi| < \delta$, in which $f'(x)$, $g'(x)$ have no simultaneous zeros. Take two points y, z in the region $|x-\xi| < \delta, \eta$. Then by the fractional formula of the mean

$$\frac{f(y)-f(z)}{g(y)-g(z)} = \frac{f'(x)}{g'(x)},$$

for some x in $)y, z($, the condition about simultaneous zeros of $f'(x)$, $g'(x)$ being satisfied.

Hence by (16)
$$\left|\frac{f(y)-f(z)}{g(y)-g(z)}-A\right| < \tfrac{1}{3}\epsilon.$$

This we may write in the form
$$\left|\frac{f(y)}{g(y)}-A-\frac{f(z)-Ag(z)}{g(y)}\right| < \tfrac{1}{3}\epsilon\left|1-\frac{g(z)}{g(y)}\right|. \tag{17}$$

We may suppose y nearer than z to ξ and consider the limit $y \to \xi$, keeping z fixed. Then $g(y) \to \infty$ and so
$$\frac{f(z)-Ag(z)}{g(y)} \to 0, \qquad \frac{g(z)}{g(y)} \to 0.$$

Hence
$$\left|\frac{f(z)-Ag(z)}{g(y)}\right| < \tfrac{1}{3}\epsilon, \qquad \left|1-\frac{g(z)}{g(y)}\right| < 2, \tag{18}$$
if $|y-\xi| <$ some $\eta'(\epsilon, z)$.

Thus from (17), (18) we have
$$|f(y)/g(y)-A| < \epsilon,$$
if $|y-\xi| < \eta, \delta, \eta'$. In other words,
$$\lim_{y\to\xi} f(y)/g(y) = A$$
and the first part of the theorem is proved. The second follows by a similar argument.

As in the case of the form $0/0$, we can discuss an indeterminate form ∞/∞ arising at infinity by making the change of arguments $x = z^{-1}$. We find at once that

(19) *If*
$$\lim_{x\to\infty} f(x) = \infty = \lim_{x\to\infty} g(x),$$
then
$$\lim_{x\to\infty} f(x)/g(x) = \lim_{x\to\infty} f'(x)/g'(x),$$
if the latter limit exists, unless $f'(x)$, $g'(x)$ have an unbounded set of common zeros.

It is evident that we can also extend the principle of (13) to the indeterminate form ∞/∞ and so assert that

(20) *In evaluating an indeterminate form ∞/∞ we may intermix the differential method of differentiating simultaneously numerator and denominator with the algebraic method of cancelling common factors, provided only that such factors do not vanish in every neighbourhood of the point of indeterminacy.*

6. Pathological phenomena

In (15) it is, of course, essential that both $f(x)$, $g(x)$ definitely diverge to infinity at ξ; a mere infinity, i.e. a point of infinite oscillation, is not enough. As an example write

$$\left.\begin{aligned} f(x) &\equiv x^{-1} \\ g(x) &\equiv x^{-1}\tan(x^{-1}) \end{aligned}\right\}, \tag{21}$$

so that at the origin $f(x)$ actually diverges, but $g(x)$ only oscillates infinitely.

Now
$$f'(x) = -x^{-2},$$
$$g'(x) = -x^{-2}\tan(x^{-1})-x^{-3}\sec^2(x^{-1}),$$

and so
$$\frac{f'(x)}{g'(x)} = \frac{x\cos^2(x^{-1})}{1+\frac{1}{2}x\sin(2x^{-1})},$$

which converges to zero at the origin.

The condition about common zeros is satisfied, since $f'(x)$ does not vanish at any finite x. Nevertheless

$$f(x)/g(x) = \cot(x^{-1}),$$

which oscillates infinitely at the origin.

The failure of (15) through violation of the condition regarding common zeros is shown by an example:†

$$\left.\begin{aligned} f(x) &\equiv 2/x+\sin(2/x) \\ g(x) &\equiv \{2/x+\sin(2/x)\}e^{\sin(1/x)} \end{aligned}\right\}. \tag{22}$$

Here $f(x)$, $g(x)$ both diverge at the origin, but

$$f(x)/g(x) = e^{-\sin(1/x)},$$

which oscillates between e and e^{-1} at the origin. Now

$$f'(x) = -2x^{-2}\{1+\cos(2/x)\},$$
$$g'(x) = -2x^{-2}\{1+\cos(2/x)+\tfrac{1}{2}[2/x+\sin(2/x)]\cos(1/x)\}e^{\sin(1/x)},$$

so that, on reduction and cancellation of the factor $\cos(1/x)$ which vanishes in every neighbourhood of the origin,

$$\frac{f'(x)}{g'(x)} = \frac{2x\cos(1/x)e^{-\sin(1/x)}}{1+x\cos(1/x)[2+\sin(1/x)]},$$

which converges to zero at the origin.

As in the case of the form 0/0, the convergence of $f(x)/g(x)$ does not imply the convergence of $f'(x)/g'(x)$. It does, however, imply, if the

† Bromwich, *Theory of Infinite Series* (1908), 381, Ex. 6 (*b*).

condition about common zeros is satisfied, that along some sequence $\{x_n\}$ of approach to ξ

$$\frac{f'(x_n)}{g'(x_n)} \to \lim_{x\to\xi} \frac{f(x)}{g(x)}.$$

As an example write

$$\left.\begin{aligned} f(x) &\equiv x^{-1}+\cos(x^{-1}) \\ g(x) &\equiv x^{-1} \end{aligned}\right\}. \tag{23}$$

Here $f(x), g(x)$ diverge at the origin and

$$f(x)/g(x) = 1+x\cos(x^{-1}),$$

which converges to 1 at the origin. But

$$f'(x) = -x^{-2}+x^{-2}\sin(x^{-1}),$$
$$g'(x) = -x^{-2},$$

so that
$$f'(x)/g'(x) = 1-\sin(x^{-1}),$$

which oscillates finitely at the origin. But

$$f'(x_n)/g'(x_n) \to 1, \quad \text{if} \quad x_n = (n\pi)^{-1}.$$

The condition about common zeros is satisfied, since $g'(x)$ does not vanish at any finite x.

If the condition about common zeros is not satisfied, it is possible, with either of the forms $0/0$, ∞/∞, for $f(x)/g(x)$, $f'(x)/g'(x)$ both to converge but to different limits. An example can be constructed (with some difficulty) to show this.

Choose any sequence $\{\xi_n\}$ converging monotonically to ξ and, when x lies in any (ξ_n, ξ_{n+1}), define

$$\left.\begin{aligned} f(x) &\equiv \frac{a_{n+1}(x-\xi_n)^2(2x+\xi_n-3\xi_{n+1})-a_n(x-\xi_{n+1})^2(2x+\xi_{n+1}-3\xi_n)}{(\xi_n-\xi_{n+1})^3} \\ g(x) &\equiv \frac{b_{n+1}(x-\xi_n)^2(2x+\xi_n-3\xi_{n+1})-b_n(x-\xi_{n+1})^2(2x+\xi_{n+1}-3\xi_n)}{(\xi_n-\xi_{n+1})^3} \end{aligned}\right\}. \tag{24}$$

Whether ξ_n be regarded as a point of either of the intervals (ξ_{n-1}, ξ_n), (ξ_n, ξ_{n+1}), the definition gives consistently

$$f(\xi_n) = a_n, \qquad g(\xi_n) = b_n.$$

The functions are therefore continuous at the points ξ_n and so also everywhere near ξ. Within an interval (ξ_n, ξ_{n+1}) we have

$$f'(x) = \frac{6(a_{n+1}-a_n)(x-\xi_n)(x-\xi_{n+1})}{(\xi_n-\xi_{n+1})^3},$$

$$g'(x) = \frac{6(b_{n+1}-b_n)(x-\xi_n)(x-\xi_{n+1})}{(\xi_n-\xi_{n+1})^3}.$$

Again, consistently, $\quad f'(\xi_n) = 0 = g'(\xi_n).$

The functions are therefore differentiable at the points ξ_n and, in fact, everywhere near ξ.

In (ξ_n, ξ_{n+1}) the function $f(x)$ has neither maximum nor minimum, since the derivative vanishes only at the end-points. The values

$$f(\xi_n) = a_n, \qquad f(\xi_{n+1}) = a_{n+1}$$

are thus the extreme values in the interval; so for $g(x)$. Hence

$$f(x), g(x) \to 0 \quad \text{at} \quad \xi,$$

if $a_n, b_n \to 0$. Moreover, in (ξ_n, ξ_{n+1}), provided that b_n, b_{n+1} have the same sign, $f(x)/g(x)$ lies between the extremes of

$$\frac{a_n}{b_n}, \quad \frac{a_{n+1}}{b_n}, \quad \frac{a_n}{b_{n+1}}, \quad \frac{a_{n+1}}{b_{n+1}}. \tag{25}$$

Finally, in (ξ_n, ξ_{n+1}),
$$\frac{f'(x)}{g'(x)} = \frac{a_n - a_{n+1}}{b_n - b_{n+1}}.$$

Now write
$$\left. \begin{aligned} a_n &= l/\sqrt{n} + (-1)^n l'/n \\ b_n &= 1/\sqrt{n} + (-1)^n/n \end{aligned} \right\}, \tag{26}$$

so that b_n is positive, if $n > 1$. Then, as $n \to \infty$, each of the fractions (25) converges to l. Hence

$$\lim_{x \to \xi} f(x)/g(x) = l.$$

But
$$a_n - a_{n+1} = \frac{l}{\{\sqrt{n}+\sqrt{(n+1)}\}\sqrt{\{n(n+1)\}}} + \frac{(-)^n(2n+1)l'}{n(n+1)},$$

$$b_n - b_{n+1} = \frac{1}{\{\sqrt{n}+\sqrt{(n+1)}\}\sqrt{\{n(n+1)\}}} + \frac{(-)^n(2n+1)}{n(n+1)},$$

and so
$$(a_n - a_{n+1})/(b_n - b_{n+1}) \to l'.$$

Thus
$$\lim_{x \to \xi} f'(x)/g'(x) = l'.$$

7. Other indeterminate forms

Certain other indeterminate forms can be reduced to one or other of the foregoing forms $0/0$, ∞/∞, and so evaluated by the methods we have been considering. They are

$$\infty - \infty, \qquad 1^\infty, \qquad 0^0, \qquad \infty^0.$$

The three exponential forms

$$1^0, \qquad 0^{+\infty}, \qquad \infty^{+\infty}$$

are, of course, not indeterminate but converge to the respective limits

$$1, \qquad 0, \qquad \infty.$$

The indeterminate form $\infty - \infty$ occurs when we seek the limit

$$\lim_{x \to \xi} \{f(x) - g(x)\},$$

given that $$\lim_{x\to\xi} f(x) = \infty = \lim_{x\to\xi} g(x).$$

We may reduce such a form to the form $0/0$ either by writing it as

$$\{[g(x)]^{-1}-[f(x)]^{-1}\}/[f(x)g(x)]^{-1}$$

or else by taking exponentials and writing

$$e^{f(x)-g(x)} = e^{-g(x)}/e^{-f(x)},$$

as appears most suitable algebraically. Each of these forms may, of course, be expressed in the form ∞/∞ and so dealt with, if that be more convenient.

In the other three indeterminate forms it is sufficient if we take logarithms. We may write symbolically

$$\log(1^\infty) = \infty.\log 1 = \infty.0,$$
$$\log(0^0) = 0.\log 0 = -0.\infty,$$
$$\log(\infty^0) = 0.\log\infty = 0.\infty.$$

In each of these the $0.\infty$ may be written in either of the forms $0/0$, ∞/∞ as is found convenient.

As elementary examples of the foregoing indeterminate forms we may consider the limits at the origin of the following functions, for convenience restricting x to be positive:

$$\operatorname{cosec} x - x^{-1} \qquad (\infty-\infty), \tag{27}$$
$$x^{-1}+\log x \qquad (\infty-\infty), \tag{28}$$
$$(1+x)^{\log x} \qquad (1^{-\infty}), \tag{29}$$
$$x^{\log(1+x)} \qquad (0^0), \tag{30}$$
$$(x^{-1})^x \qquad (\infty^0). \tag{31}$$

We may write (27) in the form $0/0$ as

$$(x-\sin x)/x\sin x,$$

or sufficiently, since $(\sin x)/x \to 1$, as

$$(x-\sin x)/x^2.$$

Two differentiations replace this by

$$(\sin x)/2,$$

which converges to the limit zero.

In (28) take exponentials and consider

$$e^{(1/x+\log x)} = e^{1/x}/x^{-1}.$$

Since $x > 0$, this has the form ∞/∞. One differentiation gives

$$-x^{-2}e^{1/x}/(-x^{-2}) = e^{1/x} \to +\infty.$$

Hence $$\lim_{x\to 0+}(x^{-1}+\log x) = +\infty.$$

In (29) we have, on taking logarithms,

$$\log(1+x)\log x.$$

Since $x^{-1}\log(1+x) \to 1$, we may write this

$$\log x/x^{-1},$$

which is of the form ∞/∞. One differentiation gives

$$-x^{-1}/x^{-2} = -x \to 0.$$

Hence
$$\lim_{x\to 0+} (1+x)^{\log x} = 1.$$

In (30), again taking logarithms, we have the same function

$$\log(1+x)\log x.$$

Hence
$$\lim_{x\to 0+} x^{\log(1+x)} = 1.$$

Finally, in (31) we have, on taking logarithms,

$$-x\log x.$$

This, as we have already seen, converges to zero. Hence

$$\lim_{x\to 0+} (x^{-1})^x = 1.$$

8. Orders of zeros and infinities

If $f(x)$, $g(x)$ have a common zero at ξ (or at infinity) and, if, as $x \to \xi$ (or $\to \infty$), the ratio $|f(x)/g(x)|$

(i) $\to 0$; (ii) has positive bounds;† (iii) $\to \infty$;

we say that the zero of $f(x)$ is of

(i) *higher order*; (ii) *equal order*; (iii) *lower order*

than the zero of $g(x)$.

If, less stringently, the ratio $|f(x)/g(x)|$

(iv) is bounded above; (v) has a positive lower bound,

so that it can oscillate to zero or to infinity respectively, we may say similarly that the zero of $f(x)$ is of order

(iv) *not higher*, (v) *not lower*

than the zero of $g(x)$. Finally, if the ratio can oscillate to both zero and infinity, we must say that the zeros of $f(x)$, $g(x)$ are *not comparable*.

In similar fashion, if $f(x)$, $g(x)$ simultaneously diverge at ξ (or at infinity), we can attribute relative orders to these infinities (or declare them incomparable). Most simply, we can take over the previous definitions by declaring that the relative order of the infinities is to be defined as the relative order of the zeros of $1/f(x)$, $1/g(x)$ at the same point.

† i.e. $A > |f(x)/g(x)| > B > 0$, so that $f(x)/g(x)$ does not oscillate either to zero or to infinity.

It follows from the two fundamental theorems of this chapter that in seeking to compare the orders of $f(x)$, $g(x)$ at a common pole we may, in general, replace them by their derivatives $f'(x)$, $g'(x)$; so too at a common zero of $f(x)$, $g(x)$ that is also a zero of $f'(x)$, $g'(x)$. It is sufficient, however, to restrict our attention to the single case of functions which diverge at infinity, since the other cases of poles or zeros can be reduced to this by taking the reciprocals of arguments or functions.

Now the successive integral powers of a variable x

$$x, x^2, \ldots, x^n, x^{n+1}, \ldots \tag{32}$$

taken at infinity, form a natural and convenient 'scale' of infinities of ascending order. The interstitial spaces of the scale will, of course, be occupied by fractional powers and, if we are still more microscopic, by irrational powers. Where in this scale shall we place the infinity $\log x$?

Now the corresponding scale of derivatives is

$$1, 2x, \ldots, nx^{n-1}, (n+1)x^n, \ldots . \tag{33}$$

The derivative of $\log x$ is x^{-1}, which precedes every member of the scale. Moreover, it precedes $\epsilon x^{\epsilon-1}$, the derivative of x^ϵ, if ϵ is any positive number no matter how small. Hence, by (15),

(34) $\log x$ *diverges with* x *more slowly than any positive power of* x.

More generally, $(\log x)^n$ diverges more slowly than any $x^{n\epsilon}$, i.e. than any $x^{\epsilon'}$. Thus the scale of infinities

$$\log x, (\log x)^2, \ldots, (\log x)^n \tag{35}$$

precedes the scale $\qquad x, x^2, \ldots, x^n$

and is, in its turn, preceded by the still more delicate scale

$$\operatorname{loglog} x, (\operatorname{loglog} x)^2, \ldots, (\operatorname{loglog} x)^n, \ldots \tag{36}$$

and so on endlessly.

Moreover, at any point x^a of the scale of powers (32) we may insert logarithmic scales either beyond the point

$$x^a(\log x)^n, \qquad x^a(\operatorname{loglog} x)^n, \ldots$$

or before the point

$$x^a(\log x)^{-n}, \qquad x^a(\operatorname{loglog} x)^{-n}, \ldots .$$

These, of course, all lie between $x^{a-\epsilon}$, $x^{a+\epsilon}$ where ϵ is positive however small.

We may similarly at any point of a logarithmic scale insert logarithmic scales of greater delicacy and insert this extended logarithmic scale in the original scale of powers.

The typical member of such an extended scale is of the form

$$x^a(\log x)^b(\log\log x)^c\ldots,$$

where the indices a, b, c,... may be any real numbers positive, negative, or zero, *so long as the first non-vanishing index be positive.*

Since x diverges more rapidly than any $(\log x)^a$, it follows, if we replace the divergent argument x by the divergent argument e^x, that

(37) *e^x diverges with x more rapidly than any power of x.*

The scale of powers of x is thus *followed* by the scale of powers of e^x

$$e^x, e^{2x},\ldots, e^{nx},\ldots. \tag{38}$$

We may replace this by the more rapidly divergent scale

$$e^x, e^{x^2},\ldots, e^{x^n}\ldots. \tag{39}$$

Still more rapidly divergent is the scale

$$\exp(e^x), \exp(e^{x^2}),\ldots, \exp(e^{x^n})\ldots, \tag{40}$$

and so on endlessly.

We may also write the simple scale of powers (32) in the form

$$e^{\log x}, e^{2\log x},\ldots, e^{n\log x},\ldots.$$

It is then seen that every scale

$$e^{(\log x)^a}, e^{2(\log x)^a},\ldots, e^{n(\log x)^a},\ldots \tag{41}$$

follows the scale of powers but precedes the exponential scales, if $a > 1$, and precedes the scale of powers but follows the logarithmic scales, if $0 < a < 1$.

In general, exponentials provide scales of increasing and unlimited power, logarithms scales of increasing and unlimited delicacy. The underlying principle is clear though a comprehensive statement is difficult.

9. The 'order' symbols. Successive approximation

If, at the origin or infinity, the functions $f(x)$, $g(x)$ have zeros or infinities of the same order, we write in a convenient notation

$$f(x) = O\,g(x). \tag{42}$$

On the other hand, if $f(x)$ has a zero of higher order or an infinity of lower order than $g(x)$, so that $f(x)/g(x) \to 0$ as $x \to 0$ or $\to \infty$ (as the case may be), I shall write†

$$f(x) = \epsilon\, g(x). \tag{43}$$

We may call these symbols O, ϵ *order* symbols.

Though we have defined, not the isolated order symbols O, ϵ them-

† I prefer this to the orthodox symbol $o\{g(x)\}$

selves, but the assertions $f(\ ) = O(\ )$, $f(\ ) = \epsilon(\ )$, and though these assertions are assertions not of equality but of inequality, yet it is still possible to establish certain laws of combination of the isolated symbols, as if they were symbols of specific functional form.

Thus, if m, n are any positive integers, it is clear that, whether $x \to 0$ or $\to \infty$, we may write

$$Ox + \epsilon x = Ox,$$
$$(Ox^m)\times(Ox^n) = Ox^{m+n},$$
$$(Ox^m)\times(\epsilon x^n) = \epsilon x^{m+n},$$
$$(Ox)^m = Ox^m,$$
$$(\epsilon x)^m = \epsilon x^m.$$

The chief use of these order symbols O, ϵ is to give an indication of the closeness of an approximation in terms of a 'small number' or 'large number' x. Thus we may write Taylor's theorem in the form

$$f(a+h) = \sum_{r=0}^{n} \frac{h^r}{r!} f_r(a) + Oh^{n+1}, \tag{44}$$

if we know that $f_{n+1}(x)$ exists; or, more exactly, in the form

$$f(a+h) = \sum_{r=0}^{n} \frac{h^r}{r!} f_r(a) + \epsilon h^n, \tag{45}$$

whether $f_{n+1}(x)$ exists or not. Here h is the 'small number' of the approximation.

If a function $y(x)$ is given implicitly by an equation of the form

$$y = a + x\phi(y), \tag{46}$$

we are accustomed, as an elementary practice, to determine an approximate solution, correct to some integral power of x, by the method of 'successive approximations'. It is worth while to consider this method here and show that it does in fact give a solution to the approximation it professes.

In this method of 'successive approximation' we take a first approximation corresponding to $x = 0$, i.e. we take

$$y_1 = a.$$

As a second approximation we substitute the first approximation for y on the right, getting

$$y_2 = a + x\phi(y_1).$$

So generally, if y_n is the nth approximation, we define the $(n+1)$th by the recurrence-formula

$$y_{n+1} = a + x\phi(y_n). \tag{47}$$

In practice we are often able, by expansion or other algebraic device,

to reject from $\phi(y_n)$ terms of order ϵx^n. We thus write our recurrence-formula (47) in the more general form

$$y_{n+1} = a+x\{\phi(y_n)+\epsilon x^n\}. \tag{48}$$

Now the method of successive approximation claims that the nth approximation y_n gives an expression for y correct to x^n, i.e. that

$$y = y_n+\epsilon x^n. \tag{49}$$

This we can prove by induction, provided that the function ϕ satisfies certain conditions. For, by (46), (48), we have

$$y = y_{n+1}+x\{\phi(y)-\phi(y_n)\}+\epsilon x^{n+1}.$$

Hence $$y = y_{n+1}+\epsilon x^{n+1},$$

if $$\phi(y)-\phi(y_n) = \epsilon x^n.$$

Since $$y-y_n = \epsilon x^n$$

by the inductive hypothesis, it is sufficient to prove that

$$\{\phi(y)-\phi(y_n)\}/(y-y_n)$$

is bounded as $x \to 0$.

If ϕ is differentiable, we may write this fraction as $\phi'(\eta)$, where η lies in the interval $)y, y_n)$. It is therefore sufficient for the induction that $\phi'(y)$ be bounded near $x = 0$.

If $\phi'(y)$ be bounded near $x = 0$, then, by chapter IV (64), $\phi(y)$ is also bounded near $x = 0$. Hence, from (46),

$$y-a \to 0, \quad \text{as } x \to 0.$$

This we may write as $$y = a+\epsilon 1,$$

which provides the first step of the induction. We may sum up the foregoing in the theorem

(50) *If the implicit relation*

$$y = a+x\,\phi(y)$$

defines a function $y(x)$ such that $\phi'(y)$ exists and is bounded near $x = 0$, then the formula of successive approximation

$$y_{n+1} = a+x\{\phi(y_n)+\epsilon x^n\}$$

defines $y_n(x)$ as an approximation to $y(x)$ correct to x^n, i.e. such that

$$y = y_n+\epsilon x^n.$$

This theorem so stated is not completely satisfactory, since we are silent about the kind of function defined by the implicit relation. This point we shall be able to take up in chapter XI.

10. Lagrange's series

It is possible, following Lagrange, to give an explicit form for the nth approximation y_n which arises from the method of successive approximation discussed in the previous section. Lagrange's formula actually goes farther and gives an expression for the nth approximation to $f(y)$ where $f(\ )$ is a function subject to certain restrictions.

We continue to take the implicit relation in the form

$$y = a + x\phi(y), \tag{51}$$

and it is convenient to anticipate the theory of chapter XI so far as to assume that

(52) 'If $\phi'(y)$ is bounded near $y = a$, then the implicit relation (51) defines a function $y(x,a)$ which is differentiable in (x,a) near $x = 0$ and has the value a at $x = 0$.'

We may state the theorem in the form:

(53) *If, for $n > 1$, $f_n(y)$, $\phi_n(y)$ exist near $y = a$, then the implicit relation*

$$y = a + x\phi(y)$$

defines a function $y(x,a)$ such that

$$f(y) = f(a) + \sum_{r=1}^{n} \frac{x^r}{r!} \frac{d^{r-1}}{da^{r-1}}\{[\phi(a)]^r f'(a)\} + \epsilon x^n.$$

Since $n > 1$ and $\phi_n(y)$ exists near $y = a$, then $\phi'(y)$ being differentiable is continuous and so bounded near $y = a$. Hence, by the assumption (52), the function $y(x,a)$ is differentiable near a. In particular we find for the partial derivatives

$$\frac{\partial y}{\partial x} = \phi(y)/\{1 - x\phi'(y)\}, \qquad \frac{\partial y}{\partial a} = 1/\{1 - x\phi'(y)\}. \tag{54}$$

It follows that $1 - x\phi'(y)$ cannot vanish near $x = 0$, $y = a$, as is otherwise obvious, since $\phi'(y)$ is bounded near $y = a$; and therefore that $\partial y/\partial x$, $\partial y/\partial a$, being differentiable functions of differentiable functions, are themselves differentiable near $y = a$. Arguing inductively in this way we can at length show that all the derivatives of y exist up to the nth order and that all the partial derivatives up to the $(n-1)$th order are differentiable. Hence, by chapter VI (29), the partial differential operators $\partial/\partial x$, $\partial/\partial a$ are commutative up to order n.

Since $f_n(y)$ exists near $y = a$, it follows that $\partial_x^n f(y)$ exists near $x = 0$, and therefore, by Taylor's formula (45), we may write

$$f(y) = f(a) + \sum_{r=1}^{n} \frac{x^r}{r!} \left\{\frac{\partial^r}{\partial x^r} f(y)\right\}_{x=0} + \epsilon x^n.$$

To evaluate the coefficients of x^r on the right observe that we have from (54)

$$\frac{\partial y}{\partial x} = \phi(y)\frac{\partial y}{\partial a}, \tag{55}$$

and that, if $g(y)$ be any differentiable function of y, then

$$\frac{\partial}{\partial x}\left\{g(y)\frac{\partial y}{\partial a}\right\} = g(y)\frac{\partial^2 y}{\partial x\partial a}+g'(y)\frac{\partial y}{\partial x}\frac{\partial y}{\partial a}$$
$$= \frac{\partial}{\partial a}\left\{g(y)\frac{\partial y}{\partial x}\right\},$$

if the operators $\partial/\partial x$, $\partial/\partial a$ are commutative. Hence near $x = 0$ we have

$$\frac{\partial}{\partial x}\left\{[\phi(y)]^r f'(y)\frac{\partial y}{\partial a}\right\} = \frac{\partial}{\partial a}\left\{[\phi(y)]^r f'(y)\frac{\partial y}{\partial x}\right\}$$
$$= \frac{\partial}{\partial a}\left\{[\phi(y)]^{r+1} f'(y)\frac{\partial y}{\partial a}\right\}, \text{ by (55)},$$

and so $$\frac{\partial}{\partial x}\frac{\partial^{r-1}}{\partial a^{r-1}}\left\{[\phi(y)]^r f'(y)\frac{\partial y}{\partial a}\right\} = \frac{\partial^r}{\partial a^r}\left\{[\phi(y)]^{r+1} f'(y)\frac{\partial y}{\partial a}\right\}.$$

Now, initially, by (55)

$$\phi(y)f'(y)\frac{\partial y}{\partial a} = f'(y)\frac{\partial y}{\partial x}.$$

Hence, by induction,

$$\frac{\partial^r}{\partial x^r}f(y) = \frac{\partial^{r-1}}{\partial a^{r-1}}\left\{[\phi(y)]^r f'(y)\frac{\partial y}{\partial a}\right\}.$$

Since x, a have been treated as independent variables, we may put $x = 0$ on the right without affecting the partial differentiations in a. This substitution gives

$$y = a, \qquad \frac{\partial y}{\partial a} = 1,$$

and so $$\left\{\frac{\partial^r}{\partial x^r}f(y)\right\}_{x=0} = \frac{\partial^{r-1}}{\partial a^{r-1}}\{[\phi(a)]^r f'(a)\},$$

and the theorem is established.

It remains to show that Lagrange's formula of approximation agrees with that given by the method of successive approximation. Suppose then that the two formulae are respectively

$$y(x) = a_0+a_1 x+\ldots+a_n x^n+\epsilon x^n$$

and $$y(x) = b_0+b_1 x+\ldots+b_n x^n+\epsilon x^n.$$

Then $$(a_0-b_0)+(a_1-b_1)x+\ldots+(a_n-b_n)x^n = \epsilon x^n,$$

i.e. $$\lim_{x\to 0}\frac{(a_0-b_0)+(a_1-b_1)x+\ldots+(a_n-b_n)x^n}{x^n}\to 0,$$

which is only possible if

$$a_0 = b_0, \quad a_1 = b_1, \quad \ldots, \quad a_n = b_n.$$

The two formulae are therefore identical.

11. Newton's method of approximation

We conclude with a method, derived from the generalized linear formula of the mean, chapter V (28), and due to Newton, for approximating to a root of an equation $f(x) = 0$. We use the particular formula corresponding to $n = 2$

$$f(a+h) = f(a)+hf'(a)+\tfrac{1}{2}h^2f''(a+\theta h). \tag{56}$$

Here $a+h$ will be taken as the root we are seeking and a as some convenient first approximation to it. The formula presupposes that $f''(x)$ exists throughout $(a, a+h)$ and hence that $f'(x)$ exists and is continuous throughout the same interval. More precisely, we stipulate that

(i) $f(a+h) = 0$,

(ii) $f(x), f'(x)$ do not vanish in $(a, a+h($, so that $f(x)$ has an invariable sign in the interval,

(iii) $f''(x)$ has throughout $(a, a+h($ this sign of $f(x)$.

Thus from (56), by (i),

$$0 = f(a)+hf'(a)+\tfrac{1}{2}h^2f''(a+\theta h),$$

i.e. $$-hf'(a)/f(a) = 1+\tfrac{1}{2}h^2f''(a+\theta h)/f(a)$$
$$> 1, \quad \text{by (iii)}.$$

Hence $-f(a)/f'(a)$ lies between $0, h$ and so $a-f(a)/f'(a)$ lies in $)a, a+h($.

Thus $$a_1 \equiv a-f(a)/f'(a)$$

is a better approximation than a to $a+h$. The stipulations (i), (ii), (iii) apply to a_1 and h_1 ($\equiv a+h-a_1$) just as much as to a and h. Hence, by a repetition of the argument, we see that

$$a_2 \equiv a_1-f(a_1)/f'(a_1)$$

is a better approximation than a_1 to $a+h$. Thus generally the recurrence-formula

$$a_{n+1} = a_n-\frac{f(a_n)}{f'(a_n)} \tag{57}$$

defines a monotonic sequence lying always in the interval $(a, a+h)$. It

is therefore bounded and so convergent. Suppose it converges to α. Proceeding to the limit† we have

$$\alpha = \alpha - f(\alpha)/f'(\alpha),$$

which gives $f(\alpha) = 0$, since $f'(x)$ is continuous and therefore bounded in $(a, a+h)$. But in (i), (ii) we have stipulated that $a+h$ be the only zero of $f(x)$ in $(a, a+h)$. Hence we must have $\alpha = a+h$, i.e. $a_n \to a+h$. Thus

(58) *If* (i) $f(\xi) = 0$, (ii) $f(x), f'(x)$ *do not vanish in the interval* $(a, \xi($, *and* (iii) $f''(x)$ *exists and has the invariable sign of* $f(x)$ *throughout* $(a, \xi($, *then the recurrence-formula*

$$a_{n+1} = a_n - \frac{f(a_n)}{f'(a_n)}$$

defines a sequence $\{a_n\}$ *which converges steadily to* ξ.

If $y = f(x)$ can be represented graphically, the result of (58) is evident from a figure. For, if the curve cross the x-axis at A and N is the foot of the ordinate from B on the curve, then (58) asserts that, if the arc AB is convex to the x-axis, then the tangent at B crosses the x-axis at a point nearer than N to A. If AB is concave to the x-axis, it is similarly to be seen that the tangent at B crosses the x-axis outside AN.

Suppose more generally that we are given that $f''(x)$ exists throughout (a, b) and has only a finite number of zeros (if any) in (a, b), and that $f(x)$ vanishes at ξ, not a zero of $f''(x)$, in the interval. Then, since the zeros of $f''(x)$ separate those of $f'(x)$ which again separate those of $f(x)$, the zeros of $f(x), f'(x), f''(x)$ in (a, b) are all isolated, and we can therefore choose a sub-interval (a_1, b_1) in which $f(x), f'(x), f''(x)$ have no zeros apart from $x = \xi$.

Since a derived function can change sign only by passing through zero, $f''(x)$ has a constant sign throughout (a_1, b_1), and $f'(x)$ has a constant sign in $(a_1, \xi($ and in $)\xi, b_1)$.

Suppose firstly that $f(a_1)$, $f(b_1)$ have opposite signs. Then, if $f(a_1)$ is that one of them which has the constant sign of $f''(x)$, the interval (a_1, ξ) satisfies the conditions of (58) and Newton's method is applicable.

On the other hand, suppose that $f(a_1)$, $f(b_1)$ have the same sign. For simplicity take $f(a_1)$, $f(b_1)$ positive and $a_1 < \xi < b_1$. Then $f'(x)$ must vanish in (a_1, b_1), for otherwise $f(x)$ would be monotonic in (a_1, b_1) and $f(\xi)$ would lie between $f(a_1), f(b_1)$. We have excluded from (a_1, b_1) every possible zero of $f'(x)$ except ξ. Hence $f'(\xi) = 0$, and $f'(x)$ is negative

† Since $f''(x)$ exists and $f'(x)$ does not vanish in the interval, $f(x)/f'(x)$ is differentiable and therefore continuous in the interval.

throughout $(a_1, \xi($ and positive throughout $)\xi, b_1)$. Thus $f'(x)$ increases in (a_1, b_1) and the constant sign of $f''(x)$ is therefore positive, i.e. $f''(x)$ has the sign of both $f(a_1), f(b_1)$. In this case each of the intervals (a_1, ξ), (ξ, b_1) satisfies the conditions of (58) and Newton's method gives a pair of monotonic sequences converging to ξ from opposite sides.

Thus, granted the existence of $f''(x)$ and the isolation of its zeros, we can approximate by Newton's method to every zero of $f(x)$ that is not also a zero of $f''(x)$.

We may include (58) under the following more general proposition:

(59) *If* (i) $\phi(x)$ *is continuous and steadily increasing in* (a_1, ξ), (ii) $\phi(x)-x$ *vanishes at* $x = \xi$ *but not elsewhere in* (a_1, ξ), (iii) $\phi(a_1)-a_1$ *has the sign of* $\xi - a_1$, *then the recurrence-formula*

$$a_{n+1} = \phi(a_n)$$

defines a monotonic sequence converging to ξ

For simplicity, suppose that $a_1 < \xi$, then, by (iii), $a_1 < \phi(a_1)$,

i.e. $$a_1 < a_2.$$

Again, by (i), $\phi(a_1) < \phi(\xi)$, since $a_1 < \xi$,

i.e. $$a_2 < \xi.$$

Lastly, by (i), $\phi(a_1) < \phi(a_2)$, since $a_1 < a_2$,

i.e. $$a_2 < \phi(a_2).$$

Thus a_2 lies in (a_1, ξ) and the conditions (i), (ii), (iii) of (59) are still satisfied, if we replace a_1 by a_2. Hence, by induction, they are satisfied, if we replace a_1 by any a_n of the sequence In other words,

$$a_n < a_{n+1} < \xi$$

for every n. The sequence $\{a_n\}$ is therefore monotonic and bounded by ξ. It thus converges to some limit α in (a_1, ξ).

Since $\phi(x)$ is continuous, we may proceed to the limit in the recurrence-formula, getting $\alpha = \phi(\alpha)$. By (ii) this gives $\alpha = \xi$ and the proof of (59) is complete.

To bring (58) under (59) write

$$\phi(x) \equiv x - \frac{f(x)}{f'(x)}$$

and therefore

$$\phi'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}.$$

Thus, if $f(x)$, $f''(x)$ have the same sign, $\phi(x)$ is monotonic. To secure (iii) of (59) write

$$0 = f(a+h) = f(a)+hf'(a+\theta h).$$

Now $f'(a+\theta h)$ has the sign of $f'(a)$, and so h has the sign of $-f(a)/f'(a)$, i.e., in the notation of (59), ξ has the sign of $\phi(a_1)-a_1$.

WORKED EXAMPLE

The functions $f(x)$, $g(x)$ are defined by the power-series

$$f(x) = x+ax^{m+1}+Ax^{2m+1}+Ox^{2m+2},$$
$$g(x) = x+bx^{m+1}+Bx^{2m+1}+Ox^{2m+2}.$$

$F(x)$, $G(x)$ are the corresponding inverse functions or, if many-valued, the continuous branches of such functions for which $F(0) = 0 = G(0)$. Show that

(i) $\displaystyle\lim_{x\to 0} \frac{fg(x)-gf(x)}{x^{3m+1}} = L,$

(ii) $\displaystyle\lim_{x\to 0} \frac{fg(x)+Fg(x)-gf(x)-gF(x)}{x^{4m+1}} = 2(2m+1)aL,$

(iii) $$\lim_{x\to 0} \frac{fg(x)+Fg(x)+fG(x)+FG(x)-gf(x)-gF(x)-Gf(x)-GF(x)}{x^{5m+1}}$$
$$= 2(10m^2+11m+2)abL,$$

where $L = m(Ab-Ba)-\frac{1}{2}m(m+1)ab(a-b)$ and $fg(x)$ is written in place of the more elaborate symbol $f\{g(x)\}$.

The power-series are to be supposed convergent in some interval containing the origin. Within this interval they are differentiable and by Taylor's formula we have

$$\begin{aligned}
fg(x)&-f(x)-g(x)+x \\
&= [g(x)-x]\{f'(x)-1\}+\tfrac{1}{2}[g(x)-x]^2f''(x)+O[g(x)-x]^3 \\
&= (bx^{m+1}+Bx^{2m+1})[(m+1)ax^m+(2m+1)Ax^{2m}]+ \\
&\qquad\qquad +\tfrac{1}{2}(bx^{m+1})^2m(m+1)ax^{m-1}+\epsilon x^{3m+1} \\
&= (m+1)abx^{2m+1}+[(2m+1)Ab+(m+1)aB+\tfrac{1}{2}m(m+1)ab^2]x^{3m+1}+\epsilon x^{3m+1}.
\end{aligned}$$

We interchange f, g and take the difference, getting

$$fg(x)-gf(x) = Lx^{3m+1}+\epsilon x^{3m+1},$$

and (i) follows.

Write $$\phi(x) = fg(x)-gf(x).$$

Since $f(x)$, $g(x) \to 0$ as $x \to 0$, we can choose x so small that $t = g(x)$, $t = f(x)$ lie within the respective intervals of convergence of $f(t)$, $g(t)$ and therefore $\phi(x)$ converges in some interval and is differentiable therein. In particular,

$$\phi'(x) = (3m+1)Lx^{3m}+\epsilon x^{3m}, \tag{1}$$

and generally $$\phi_r(x) = Ox^{3m-r+1} \qquad (r < 3m+1).$$

The functions f, F are connected by the relations

$$y = f(x), \qquad F(y) = x,$$

and hence $$fF(t) = t = Ff(t).$$

Thus $$x = F(x)+a\{F(x)\}^{m+1}+\epsilon\{F(x)\}^{m+1},$$

i.e. $$x/F(x) = 1+a\{F(x)\}^m+\epsilon\{F(x)\}^m.$$

But $F(x)$ is continuous and $F(0) = 0$, and we therefore have

$$\lim_{x\to 0} x/F(x) = 1. \tag{2}$$

It follows too that $$x-F(x) = ax^{m+1}+\epsilon x^{m+1}. \tag{3}$$

Again, $$F'(x) = \lim_{y\to x}\frac{F(y)-F(x)}{y-x}$$

$$= \lim_{F(y)\to F(x)}\frac{F(y)-F(x)}{fF(y)-fF(x)}, \quad \text{since } F(x) \text{ is continuous,}$$

i.e. $$F'(x) = \frac{1}{f'F(x)}. \tag{4}$$

Hence $F'(x)$ exists and in particular $F'(0) = 1$. Similarly $F''(x)$ exists, since $f''(x)$ exists. Now, if in (i) we put $F(x)$ for x, we have

$$fgF(x)-g(x) = L\{F(x)\}^{3m+1}+\epsilon\{F(x)\}^{3m+1}$$
$$= Lx^{3m+1}+\epsilon x^{3m+1}, \quad \text{by (2).}$$

Again

$$gF(x)-Fg(x) = FfgF(x)-Fg(x)$$
$$= [fgF(x)-g(x)]F'g(x)+O[fgF(x)-g(x)]^2,$$
by Taylor's theorem,
$$= [fgF(x)-g(x)]F'g(x)+Ox^{6m+2}, \quad \text{by (i),}$$
$$= \phi F(x)F'g(x)+Ox^{6m+2}.$$

Hence

$$fg(x)+Fg(x)-gf(x)-gF(x)$$
$$= \phi(x)-\phi F(x)F'g(x)+Ox^{6m+2}$$
$$= \phi(x)-\phi F(x)+\phi F(x)\{1-F'g(x)\}+Ox^{6m+2} \tag{5}$$
$$= \{x-F(x)\}\phi'(x)+\phi F(x)\{1-F'g(x)\}+Ox^{5m+1},$$

since $$x-F(x) = Ox^{m+1}, \qquad \phi_r(x) = Ox^{3m-r+1}.$$

Now $$f'(x) = 1+(m+1)ax^m+\epsilon x^m,$$

and so $$f'F(x) = 1+(m+1)ax^m+\epsilon x^m.$$

Hence, by (4), $$F'(x) = 1-(m+1)ax^m+\epsilon x^m,$$

and so $$F'g(x) = 1-(m+1)ax^m+\epsilon x^m.$$

We thus have

$$fg(x)+Fg(x)-gf(x)-gF(x)$$
$$= ax^{m+1}(3m+1)Lx^{3m}+Lx^{3m+1}(m+1)ax^m+\epsilon x^{4m+1}$$
$$= 2(2m+1)aLx^{4m+1}+\epsilon x^{4m+1},$$

which gives (ii).

Write now $$\psi(x) = fG(x)-Gf(x).$$

Then, if we interchange f, g in (ii), we have

$$\phi(x)+\psi(x) = (4m+2)bLx^{4m+1}+\epsilon x^{4m+1},$$

and consequently

$$\phi'(x)+\psi'(x) = (4m+2)(4m+1)bLx^{4m}+\epsilon x^{4m},$$

and generally $\qquad \phi_r(x)+\psi_r(x) = Ox^{4m-r+1} \qquad (r < 4m+1).$

Now, as in (5),

$$fg(x)-gf(x)+Fg(x)-gF(x)+fG(x)-Gf(x)+FG(x)-GF(x)$$
$$= \phi(x)+\psi(x)-\phi F(x)-\psi F(x)+\phi F(x)\{1-F'g(x)\}+\psi F(x)\{1-F'G(x)\}+Ox^{6m+2}$$
$$= [x-F(x)]\{\phi'(x)+\psi'(x)\}+Ox^{6m+1}+\{\phi F(x)+\psi F(x)\}\{1-F'G(x)\}+$$
$$+\phi F(x)\{F'G(x)-F'g(x)\}, \quad \text{by Taylor's theorem.}$$

But
$$x-F(x) = ax^{m+1}+\epsilon x^{m+1},$$
$$\phi'(x)+\psi'(x) = (4m+2)(4m+1)bLx^{4m}+\epsilon x^{4m},$$
$$\phi F(x)+\psi F(x) = (4m+2)bLx^{4m+1}+\epsilon x^{4m+1},$$
$$1-F'G(x) = (m+1)ax^m+\epsilon x^m,$$
$$\phi F(x) = Lx^{3m+1}+\epsilon x^{3m+1}.$$

Again
$$G(x) = x-bx^{m+1}+\epsilon x^{m+1},$$
and so
$$G(x)-g(x) = -2bx^{m+1}.$$
Hence
$$F'G(x)-F'g(x) = -2bx^{m+1}F''(x)+Ox^{2m+2}$$
$$= 2abm(m+1)x^{2m}+\epsilon x^{2m}.$$

Thus $\qquad \phi F(x)\{F'G(x)-F'g(x)\} = 2m(m+1)abLx^{5m+1}+\epsilon x^{5m+1},$

and so finally

$$fg(x)-gf(x)+Fg(x)-gF(x)+fG(x)-Gf(x)+FG(x)-GF(x)$$
$$= \{(4m+2)(4m+1)+(4m+2)(m+1)+2m(m+1)\}abLx^{5m+1}+\epsilon x^{5m+1}$$
$$= 2(11m^2+10m+2)abLx^{5m+1}+\epsilon x^{5m+1},$$

which gives (iii).

EXAMPLES VII

1. Prove that Theorem (6) remains true when all the derivatives which occur are only umbral derivatives.

2. If the conditions of Theorem (6) apply, namely that

$$f_r(\xi) = 0 \qquad g_r(\xi) \quad (r = 0,\ldots,n-1), \qquad g_n(\xi) \neq 0,$$

and if, in addition, $f_{n+1}(\xi)$, $g_{n+1}(\xi)$ exist, prove that

$$\left\{D\frac{f(x)}{g(x)}\right\}_{x=\xi} = \frac{1}{n+1}\left\{D\frac{f_n(x)}{g_n(x)}\right\}_{x=\xi}.$$

If $f_{n+2}(\xi)$, $g_{n+2}(\xi)$ exist, obtain the corresponding expression for $D^2[f(x)/g(x)]$ at $x = \xi$.

3. Evaluate the limits

$$\lim_{x\to 1}\frac{\log\{x+\sqrt{(x^2-1)}\}}{\sqrt{(x^2-1)}}, \qquad \lim_{a\to b}\frac{a^b-b^a}{a^a-b^b},$$

$$\lim_{x\to 0}\log(1-x)\cot\tfrac{1}{2}\pi x, \qquad \lim_{x\to 1}\log(1-x)\cot\tfrac{1}{2}\pi x,$$

$$\lim_{x\to 0}\left(\frac{a_1^x+\ldots+a_n^x}{n}\right)^{1/x}, \qquad \lim_{x\to 0}(1-\log|x|)^x.$$

4. Show that, as $x \to 0$,

$$C_n = \frac{1}{(n!)^2} + 2\sum_{r=1}^{n} \frac{(-)^r \cos rx}{(n+r)!(n-r)!} = Ox^{2n},$$

$$S_n = 2\sum_{r=1}^{n} \frac{(-)^r r \sin rx}{(n+r)!(n-r)!} = Ox^{2n-1},$$

$$C_n' = 2\sum_{r=1}^{n} \frac{(-)^r r^2 \cos rx}{(n+r)!(n-r)!} = Ox^{2n-2},$$

and obtain the limits, as $x \to 0$, of

$$C_n/x^{2n}, \qquad S_n/x^{2n-1}, \qquad C_n'/x^{2n-2}.$$

5. Show that

$$\lim_{x \to 0} \frac{1-\sin\{\frac{1}{2}\pi \sin[\frac{1}{2}\pi \sin \ldots \sin(\frac{1}{2}\pi \cos \frac{1}{2}\pi x)]\}}{x^N} = (\tfrac{1}{8}\pi^2)^{N-1},$$

where 'sin' occurs n times in the functional symbol and $N = 2^{n+1}$.

6. Obtain the limit $$\lim_{x \to +\infty} \frac{\log\{\log[\ldots \log(x+1)]\}}{\log\{\log[\ldots \log x]\}},$$

where 'log' occurs n *times in each functional symbol.*

7. (i) If $u_1(x) = x$, $u_2(x) = x^x$, $u_3(x) = x^{x^x}$, and generally $u_{n+1}(x) = x^{u_n(x)}$, show that, as $x \to 0+$,

$$u_{2n}(x) \to 1, \qquad u_{2n+1}(x) \to 0.$$

(ii) If $v_1(x) = e^{-x}$, $v_2(x) = \exp(-xe^{-x})$, and generally $v_{n+1}(x) = \exp[-xv_n(x)]$, show that, as $x \to +\infty$,

$$v_{2n}(x) \to 1, \qquad v_{2n+1}(x) \to 0.$$

8. Determine the limits, as $x \to \infty$, of

$$x\left\{\left(1+\frac{a}{x}\right)^x - e^a\right\}, \qquad x\left\{\left(1-\frac{a}{x}\right)^{-x} - \left(1+\frac{a}{x}\right)^x\right\}, \qquad \frac{x^{x+1}}{(x-1)^x} - \frac{(x+1)^x}{x^{x-1}}.$$

9. If $f(x) \to 0$ and $f(x)g(x) \to l$ as $x \to 0$, show that

$$\lim_{x \to 0} \{1+f(x)\}^{g(x)} = e^l.$$

If $f(x) \to 0$ and $g(x)/f(x) \to l$ as $x \to 0$, show that

$$\lim_{x \to 0} |f(x)|^{g(x)} = 1.$$

Show that $$\lim_{x \to 0} |x|^{x+\epsilon} = 0, 1, \infty$$

according as $\epsilon > 0, = 0, < 0$.

10. Obtain the limits

$$\lim_{x \to 0} (\sec^2 ax)^{\cot^2 bx}, \qquad \lim_{x \to 0} \left(\sin^2 \frac{\pi}{2-ax}\right)^{\sec^2 \frac{\pi}{2-bx}}.$$

11. If $f(x) \to 0$, $g(x) \to 0$, $\phi'(x) \to l$ as $x \to 0$, prove that

$$\frac{\phi\{f(x)\} - \phi\{g(x)\}}{f(x)-g(x)} \to l,$$

$$\frac{\{1+f(x)\}^{1/f(x)} - \{1+g(x)\}^{1/g(x)}}{f(x)-g(x)} \to -\tfrac{1}{2}e,$$

$$\frac{|f(x)|^{1+f(x)} - |g(x)|^{1+g(x)}}{f(x)-g(x)} \to 1,$$

and evaluate the limit

$$\lim_{x\to 0}\frac{|x\sin(1/x)|^{1+x\sin(1/x)}-|x\cos(1/x)|^{1+x\cos(1/x)}}{x\{\sin(1/x)-\cos(1/x)\}}.$$

12. Evaluate the limits

$$\lim_{x\to 0}\{(\cot x)^{\frac{1}{2}}\sin(\cot x)-(\operatorname{cosec} x)^{\frac{1}{2}}\sin(\operatorname{cosec} x)\},$$

$$\lim_{x\to 0+}\frac{\sin(\log x)-\sin\{\log(\sin x)\}}{x\sin x\log(x\sin x)},$$

$$\lim_{x\to\infty}\frac{\sin^2 x\sin(x\sec x)-\cos^2 x\sin(x\operatorname{cosec} x)}{x^2(\sin x-\cos x)}.$$

13. Obtain the limits, as $x \to 0+$, of the expressions

$$\frac{x^x-(\sin x)^x}{x^3},\qquad \frac{(\sin x)^x+(\sin^{-1}x)^x-2x^x}{x^5},$$

$$\frac{x^{\sin x}-(\sin x)^x}{(\sin x)^{\sin x}-x^x},\qquad \frac{x^{\sin x}-(\sin x)^{\sin x}}{(\sin x)^x-(\sin x)^{\sin x}},$$

$$\frac{(\tan x)^{\tan x}+(\tan^{-1}x)^{\tan^{-1}x}-(\sin x)^{\sin x}-(\sin^{-1}x)^{\sin^{-1}x}}{x^5\log x}.$$

14. Prove that

$$\lim_{x\to 0}\frac{\sec(\cosh x-1)-\cosh(\sec x-1)+\sec x-\cosh x}{x^8}=\frac{1}{1440}.$$

15. Prove that, as $x\to 0$,

$$\frac{\tan(\sin x)-\sin(\tan x)}{x^7}\to\frac{1}{30},$$

$$\frac{\tan(\sin x)+\tan^{-1}(\sin x)-\sin(\tan x)-\sin(\tan^{-1}x)}{x^9}\to\frac{1}{9},$$

$$\frac{\tan(\sin x)-\tan^{-1}(\sin^{-1}x)-\sin(\tan x)+\sin^{-1}(\tan^{-1}x)}{x^9}\to\frac{1}{18},$$

$$x^{-11}\{\tan(\sin x)+\tan^{-1}(\sin x)-\tan(\sin^{-1}x)+\tan^{-1}(\sin^{-1}x)-$$
$$-\sin(\tan x)-\sin(\tan^{-1}x)-\sin^{-1}(\tan x)-\sin^{-1}(\tan^{-1}x)\}\to-\frac{32}{135}.$$

16. Prove that, as $x\to 0$,

$$\frac{\log(1+x)\log(1-x)-\log(1-x^2)}{x^4}\to\frac{1}{12},$$

$$x^{-5}\left\{\left[\log\frac{1-x^2}{1+2x}\right][\log(1+x)]^2-\log\frac{(1-x)(1+x)^3}{1+2x}\right\}\to\frac{1}{2},$$

$$x^{-6}\left\{\frac{2\log(1+x)}{1-[\log(1+x)]^2}-\log\frac{1-x^2}{1-2x}-\frac{2(e^x-3)}{e^x-2}-e^{\frac{x}{1-x}}-e^{\frac{x}{1-x}}\right\}\to\frac{23}{12}.$$

17. If the necessary derivatives exist, prove that, as $x\to a$,

$$\frac{f'(x)}{f(x)-f(a)}-\frac{1}{x-a}\to\frac{1}{2}\frac{f''(a)}{f'(a)},\qquad \left\{\frac{f'(x)+f'(a)}{2f(x)-2f(a)}\right\}^2-\frac{1}{(x-a)^2}\to\frac{1}{6}\frac{f_3(a)}{f'(a)},$$

$$\left\{\frac{f'(x)+3f'(\frac{1}{3}x+\frac{2}{3}a)}{4f(x)-4f(a)}\right\}^3-\frac{1}{(x-a)^3}\to\frac{1}{72}\frac{f_4(a)}{f'(a)},$$

$$\left\{\frac{f'(x)+4f'(\frac{1}{2}x+\frac{1}{2}a)+f'(a)}{6f(x)-6f(a)}\right\}^4-\frac{1}{(x-a)^4}\to\frac{1}{720}\frac{f\,(a)}{f'(a)},$$

$$\left\{\frac{f'(x)+3f'(\frac{2}{3}x+\frac{1}{3}a)+3f'(\frac{1}{3}x+\frac{2}{3}a)+f'(a)}{8f(x)-8f(a)}\right\}^4-\frac{1}{(x-a)^4}\to\frac{1}{1620}\frac{f_5(a)}{f'(a)}.$$

18. If
$$F(x) \equiv \frac{f(x)-f(a)}{x-a}$$
and if $f_{n+1}(a)$ exists, show that
$$\lim_{x\to a} F_n(x) = \frac{f_{n+1}(a)}{n+1} = F_n(a).$$
If
$$F(x,y) \equiv \frac{f(x)-f(y)}{x-y},$$
and if $f_{n+1}(x) \to f_{n+1}(a)$ as $x \to a$, show that
$$\lim_{x,y\to a} \frac{\partial^n F(x,y)}{\partial x^n} = \frac{f_{n+1}(a)}{n+1}.$$

19. If
$$F(x,y,z) \equiv \begin{vmatrix} f(x) & f(y) & f(z) \\ g(x) & g(y) & g(z) \\ h(x) & h(y) & h(z) \end{vmatrix} \div (x-y)(x-z)(y-z),$$
show that (i) $F(x,y,a) \to \dfrac{1}{2!}(D,1)^2|f(a)\ g(a)\ h(a)|$,

as $x, y \to a$, if $f''(a)$, $g''(a)$, $h''(a)$ exist;

(ii) $F(x,y,z) \to \dfrac{1}{2!}(D,1)^2|f(a)\ g(a)\ h(a)|$,

as $x,y,z \to a$, if $f''(x)$, $g''(x)$, $h''(x)$ are continuous at $x = a$.

20. If $f(x) \to \infty$ as $x \to \infty$, show that
$$\lim_{x\to\infty} f'(x) = l$$
secures
$$\lim_{x\to\infty} \{f(x+1)-f(x)\} = l,$$
which again secures
$$\lim_{x\to\infty} f(x)/x = l,$$
but that we cannot argue in the reverse direction, even if $f(x)$ be monotonic.

21. If, as $x \to \infty$, $\quad f'(x) = $ (i) Ox^r, $\quad$ (ii) ϵx^r $\quad (r > 0)$,

show that respectively $\quad f(x) = $ (i) Ox^{r+1}, $\quad$ (ii) ϵx^{r+1}.

If $f(0) = 0$ and if, as $x \to 0$,

$f'(x) = $ (i) Ox^r, $\quad$ (ii) ϵx^r $\quad (r > 0)$,

show that respectively $\quad f(x) = $ (i) Ox^{r+1}, $\quad$ (ii) ϵx^{r+1}.

Show, however, that we cannot argue conversely from $f(x)$ to $f'(x)$.

22. If $\theta(a,h)$ be defined by the formula of the mean
$$f(a+h) = f(a) + hf'(a+\theta h),$$
show that, if $f_4(a)$ exists and $f_2(a) \neq 0$, then
$$\theta(a,h) = \frac{1}{2} + \frac{h}{24}\frac{f_3(a)}{f_2(a)} + \frac{h^2}{48}\frac{f_2(a)f_4(a)-f_3^2(a)}{f_2^2(a)} + \epsilon h^2.$$

23. If near $x = 0$, $y = 0$, we have
$$y(x) = x + a_2 x + \dots + a_n x^n + \epsilon x^n$$
and, for the corresponding inverse function,
$$x(y) = y + b_2 y + \dots + b_n y^n + \epsilon y^n,$$
show that, for $r \leqslant n$,
$$b_r = (-)^{r-1}\{D^2, D^3, \dots, D^r\}|y, y^2, \dots, y^{r-1}|_{x=0},$$
the constituents of the sth row being the coefficients of x^{s+1} in $y, y^2, \dots, y^{r-1}$.

24. If $y = a+x\phi(y)$ and

$$X_1 = \frac{x\phi(a)}{1-x\phi'(a)}, \qquad X_2 = \frac{x\phi(a)}{1-x\phi'(y)},$$

show that, granted differentiability,

$$\lim_{x\to 0}\frac{y-a-X_1}{X_1^2} = \frac{\phi''(a)}{2\phi(a)} = -\lim_{x\to 0}\frac{y-a-X_2}{X_2^2},$$

$$\lim_{x\to 0}\frac{2y-2a-X_1-X_2}{X_1^2 X_2^2} = \frac{\phi_3(a)}{6\phi(a)}.$$

25. If $y = a+x\phi(y)$, show that the operators

$$\frac{d}{dx}, \qquad \frac{1}{1-x\phi'(y)}\frac{\partial}{\partial y},$$

operating on any $\psi(x, y)$, are commutative. Show also that

$$\frac{d^n f(y)}{dx^n} = \left[\frac{1}{1-x\phi'(y)}\frac{\partial}{\partial y}\right]^{n-1}\left\{\frac{f'(y)[\phi(y)]^n}{1-x\phi'(y)}\right\}.$$

26. Obtain the Lagrange's expansions of

(i) $e^{ay}\cos by$ in powers of ye^{cy},

(ii) $\log(ay+b)$ in powers of $y(ay+b)^r$,

(iii) y in powers of $y^{-1}\log y$.

27. Apply Lagrange's method to obtain

(i) the expansion of $\log\{1+\sqrt{(1-x)}\}$ in ascending powers of x,

(ii) a series for $y-a$ in fractional descending powers of $a-b$, where

$$(y-a)^p(y-b)^q = 1.$$

28. If, under the conditions of Lagrange's theorem,

$$y = a+x\phi(y),$$

show that

$$\text{(i)}\quad \frac{1}{1-x\phi'(y)} = \sum_{r=0}^{n}\frac{x^r}{r!}\frac{d^r}{da^r}\{\phi(a)\}^r + \epsilon x^n,$$

$$\text{(ii)}\quad \frac{(y-a)^k}{k} = \sum_{r=0}^{n}\frac{x^{k+r}}{r!(r+k)}\frac{d^r}{da^r}\{\phi(a)\}^{r+k} + \epsilon x^{n+k}.$$

Show that

$$(1-2ax+x^2)^{-\frac{1}{2}} = \sum_{r=0}^{n}\frac{x^r}{r!}\frac{d^r}{da^r}\left(\frac{a^2-1}{2}\right)^r + \epsilon x^n.$$

29. If $z = a+x\phi(z)+y\psi(z)$, show that, granted differentiability,

$$\frac{\partial^{p+q}f(z)}{\partial x^p\partial y^q} = \left[\frac{1}{1-x\phi'(z)-y\psi'(z)}\frac{\partial}{\partial z}\right]^{p+q-1}\frac{f'(z)[\phi(z)]^p[\psi(z)]^q}{1-x\phi'(z)-y\psi'(z)},$$

where, on the right, x, y are to be regarded as constants during the differentiations in z.

Show also that z satisfies the partial differential equation

$$\left(\frac{\partial z}{\partial y}\right)^2\frac{\partial^2 z}{\partial x^2} - 2\frac{\partial z}{\partial x}\frac{\partial z}{\partial y}\frac{\partial^2 z}{\partial x\partial y} + \left(\frac{\partial z}{\partial x}\right)^2\frac{\partial^2 z}{\partial y^2} = 0.$$

Generalize these results for an implicit function of n independent variables.

30. If $z = a+x\phi(z)+y\psi(z)$, show that, granted differentiability,

$$f(z) = f(a)+\sum_{p,q=1}^{p+q=n} \frac{x^p y^q}{p!\,q!}\left(\frac{\partial}{\partial a}\right)^{p+q-1}\{f'(a)[\phi(a)]^p[\psi(a)]^q\}+\sum_{r=0}^{n} x^r y^{n-r}\epsilon_r,$$

where $\epsilon_0,\ldots,\epsilon_n \to 0$ as $x, y \to 0$.

31. Obtain the sequences given by Newton's method for approximating to zeros of the functions x^a, $(x-a)^{b/(b-a)}(x-b)^{a/(a-b)}$,

and determine the necessary relations between the constants in order that the conditions of the method may be satisfied.

32. The sequences $\{x_n\}$, $\{x_n'\}$ are obtained by applying Newton's method of approximation to determine $\sqrt{a}$ as a zero of the respective functions x^2-a, $x-a/x$. Prove that, if $x_0 x_0' = a$, then, for every n,

$$x_n x_n' = a.$$

33. Discuss the convergence of the sequences defined by the following recurrence-formulae:

(i) $a_{n+1} = \log(1+a_n)$, $\quad |a_0| < 1$,

(ii) $a_{n+1} = \sin \frac{1}{2}\pi a_n$,

(iii) $2a_{n+1} = a_n^2+c^2+1$,

(iv) $2a_{n+1} = a_n^2-c^2+1$, $\quad a_0 > 1+|c|$,

(v) $2a_{n+1} = a_n^2-c^2+1$, $\quad |a_0-1| < |c| < 1$.

34. If Newton's method of approximation be applied to a function $f(x)$ in which $f(x)$, $f''(x)$ are always positive (and never zero) for real values of x, show that the sequence

$$a_{n+1} = a_n - \frac{f(a_n)}{f'(a_n)}$$

diverges to infinity, if $f'(x)$ has no real zero, or oscillates, if $f'(x)$ has a real zero, unless, exceptionally, some a_n actually attains to this zero.

Discuss the examples

(i) $f(x) = x^2-x+1$, $\quad$ (ii) $f(x) = e^x$.

VIII

ANALYTIC FUNCTIONS

1. Expansion in power-series

In our original definition of functional dependence we were careful to distinguish the form of the functional relation from the analytical expressions (if any) which might be used to represent it. From the theoretical point of view the fact of functional dependence and the form of the functional relation are chiefly important, but it is clear that we shall be handicapped for practical purposes, if we cannot find convenient analytical representation of the functional relation.

By 'analytical expression' we think more particularly of functions expressed as the result of algebraic operations combined, if necessary, with passage to the limit. The typical such expression is the limit-function of a sequence of elementary functions. More precisely, it is as the sum-function of a series of elementary functions that we find our most fertile mode of analytical expression. Simplest among such series of elementary functions is the *power-series*, which we discuss in this chapter.

The form of Taylor's series suggests that representation by power-series is peculiarly relevant to the methods and ideas of the Differential Calculus, and in the preceding chapter we have already seen the possibility of approximating to a function by a terminated series in powers of x, the approximation being correct to some specified power of x. We have now, however, to consider the more exacting problem of representing a given function, not approximately but precisely, by a power-series, the representation being valid over some prescribed domain.

We shall first show that certain conditions must be satisfied by a function $f(x)$, if it can be represented by a power-series, and then pass on to prove that these conditions are also sufficient to ensure the representation.

In general terms we may describe the class of functions which can be represented by power-series as the class of *analytic* functions. We shall convert this loose description into a precise definition at a later stage. Meanwhile we prepare the way for the discussion of this wider class of functions by considering the conditions which characterize a subdivision of this class, namely the class of sum-functions of power-series.

2. Properties of the sum-function of a power-series

The properties of a series of powers of $x+a$ will clearly not differ essentially from those of a series of powers of x, and it is therefore sufficient to consider only series of the latter type. Write

$$f(x) \equiv \sum_{r=0}^{\infty} a_r x^r. \tag{1}$$

The series has an interval of convergence $(-\rho, \rho)$, where

$$\rho = \varliminf |a_n|^{-1/n}. \tag{2}$$

If $\rho = 0$, $f(x)$ is defined only at the point $x = 0$ and not over an interval. Hence we suppose

$$\rho > 0. \tag{3}$$

The interval of convergence may be open, half-open, or closed. Within this interval, by chapter IV (69), the sum-function is differentiable and its derivative is represented by the differentiated series, thus

$$f'(x) = \sum_{r=1}^{\infty} r a_r x^{r-1}. \tag{4}$$

We may apply chapter IV (69) to this new series and so on, until at length, for any n,

$$f_n(x) = \sum_{r=0}^{\infty} \frac{(n+r)!}{r!} a_{n+r} x^r. \tag{5}$$

Thus

(6) *The sum-function of a power-series has derivatives of every order throughout the open interval of convergence.*

We are required to say 'open' interval, since the theorem may or may not extend to the end-points of convergence. Thus

$$f(x) \equiv \sum_{r=1}^{\infty} \frac{x^{2r}+x^{2r+1}}{r(r+1)} \tag{7}$$

has the closed interval of convergence $(-1, 1)$. The first derivative exists at -1, but not at $+1$; the second derivative exists at neither end-point. These facts become evident, if we sum the series and write

$$f(x) = \frac{(1-x)(1+x)^2 \log(1-x^2)}{x^2} - (1+x).$$

On the other hand write†

$$f(x) \equiv \sum_{r=1}^{\infty} e^{-\sqrt{r}} x^r. \tag{8}$$

Then

$$a_r^{1/r} \to 1,$$

† I owe the example to Professor G. H. Hardy.

and so the interval of convergence is $(-1, 1)$. Again by the 'logarithmic' tests for convergence† it is seen that, for any n,

$$\sum \frac{(n+r)!}{r!} e^{-\sqrt{r}}$$

converges. Hence derivatives of every order exist throughout the closed interval of convergence.

Now (5) holds throughout the interval $)-\rho, \rho($ and therefore in particular at $x = 0$. Substitution gives

$$a_n = f_n(0)/n!, \tag{9}$$

and we may therefore rewrite (1) in the form

$$f(x) = \sum_{r=0}^{\infty} \frac{x^r}{r!} f_r(0), \tag{10}$$

which is exactly the Taylor's series for $f(x)$ with $x = 0$ as 'point of origin'.‡ Thus

(11) *A power-series is a Taylor's series of its own sum-function, the point of origin being the mid-point of the interval of convergence.*

From (2), (9) we have

$$\rho = \lim |f_n(0)/n!|^{-1/n},$$

and hence supposing, as we do, that $\rho > 0$ we must have that the sequence

$$|f_n(0)/n!|^{1/n}$$

is bounded. We can go farther and prove that

(12) *If $f(x)$ is the sum-function of a power-series convergent in $(-\rho, \rho)$, then*

$$|f_n(x)/n!|^{1/n}$$

is bounded, when n is a positive integer and x lies in $(-\sigma, \sigma)$ where $\sigma < \rho$.

Choose any σ_1 between σ, ρ, so that

$$\sigma < \sigma_1 < \rho,$$

and, in virtue of (2),

$$|a_n|^{-1/n} > \sigma_1, \quad \text{if } n > \text{some } N(\sigma_1),$$

i.e.

$$|a_n| < \sigma_1^{-n}.$$

† Cf. Bromwich, *Theory of Infinite Series* (1908), § 11.

‡ The Taylor's series

$$f(x) = \sum_{r=0}^{\infty} \frac{(x-a)^r}{r!} f_r(a)$$

is, of course, a function of a as well as of $f(x)$, and it is convenient to bring a into the title of the series under some such phrase as 'the point of origin'. The particular series (8) formed with $x = 0$ as point of origin is sometimes distinguished as 'Maclaurin's series'. There seems, however, no real need for the different name, especially since it does not appear that Maclaurin imagined his series to be anything more than a particular case of Taylor's series.

Hence, from (5), $|f_n(x)/n!| < \sum_{r=0}^{\infty} \frac{(n+r)!}{n!\,r!} |x|^r \sigma_1^{-(n+r)}$,

i.e. $< \sigma_1^{-n}\{1-|x|\sigma_1^{-1}\}^{-(n+1)}$, since $|x| \leqslant \sigma < \sigma_1$.

Thus we have
$$|f_n(x)/n!| < \sigma_1(\sigma_1-|x|)^{-(n+1)}, \tag{13}$$
which we shall need subsequently.

Widening the inequality we may write†
$$|f_n(x)/n!| < \sigma_1^n(\sigma_1-|x|)^{-2n},$$
i.e.
$$|f_n(x)/n!|^{1/n} < \sigma_1(\sigma_1-|x|)^{-2} < \sigma_1(\sigma_1-\sigma)^{-2},$$
and so is bounded, if $n > N$. But each of the functions
$$|f_n(x)/n!|^{1/n} \qquad (n = 1, 2,..., N)$$
is continuous and so bounded in $(-\sigma, \sigma)$. Thus we may remove the restriction $n > N$, and (12) is therefore established.

3. Behaviour at the boundary of convergence

If the interval of convergence is open it is not to be expected that in general
$$F(n, x) \equiv |f_n(x)/n!|^{1/n} \tag{14}$$
will be bounded in the full open interval of convergence. This can be confirmed by an example. For write
$$f(x) \equiv \sum_{r=0}^{\infty} x^r.$$
Then
$$|f_n(x)/n!|^{1/n} = (1-x)^{-(1+1/n)},$$
which is unbounded in the interval of convergence $)-1, 1($.

Suppose now that the interval of convergence is closed, at any rate at one end, say $x = \rho$. Then there are three alternative possibilities:

(i) every differentiated series also converges at ρ and therefore defines at ρ the corresponding derivative $f_n(\rho)$;

(ii) the differentiated series do not all converge at ρ, but nevertheless every derivative $f_n(x)$ exists at $x = \rho$;

(iii) the derivatives of $f(x)$ do not all exist at $x = \rho$. Since the existence of $f_n(\rho)$ implies the existence of every preceding derivative at and near ρ, this means in effect that $f_n(\rho)$ does not exist after some specifiable value of n.

In cases (i), (ii) it may or it may not happen that $F(n, x)$ is bounded in the closed interval $(0, \rho)$. In case (iii) the question does not properly arise, since $F(n, \rho)$ ceases to exist after some definite value of n.

† Since $\sigma_1^{n-1} > (\sigma_1-|x|)^{n-1}$.

We justify these assertions in case (i) by referring to the worked example at the end of this chapter.† There we define a function $f(x)$ such that every $f_n(x)$ exists throughout the closed interval of convergence $(-1, 1)$; such that $F(n, x)$ is bounded in any interval $(-1, \sigma)$ where $0 < \sigma < 1$; and such that $F(n, x)$ is unbounded in the full interval of convergence $(-1, 1)$. Hence, in particular $f(x)$ itself is a function such that $F(n, x)$ is unbounded in $(0, 1)$, while $f(-x)$ is a function such that $F(n, x)$ is bounded in $(0, 1)$, both the functions belonging to case (i).

For case (ii) consider the function

$$\bar{f}(x) = x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \ldots.$$

Here $\rho = 1$ and none of the differentiated series converge at $x = \rho$. Nevertheless, every $\bar{f}_n(\rho)$ exists and in fact, if $0 < x \leqslant 1$,

$$\bar{f}_n(x) = (-)^{n-1}(n-1)!\,(1+x)^{-n}.$$

Thus $\bar{F}(n, x)$ is bounded in $(0, 1)$.

For an example to the contrary we define the function $f(x)+\bar{f}(x)$, where $f(x)$, $\bar{f}(x)$ are the functions just considered, by adding together the corresponding power-series. The combined power-series has the same ρ of convergence. The differentiated series do not converge at ρ, since those for $f(x)$ do but those for $\bar{f}(x)$ do not. But every derivative of the combined function exists at ρ, since every $f_n(\rho)$, $\bar{f}_n(\rho)$ separately exist. Thus the example falls under case (ii), and we have

$$\begin{aligned} F(n, x) &= |f_n(x)/n! + \bar{f}_n(x)/n!|^{1/n} \\ &\geqslant |f_n(x)/n!|^{1/n} - |\bar{f}_n(x)/n!|^{1/n}. \end{aligned}$$

This is the difference of two expressions the first of which is unbounded, the second bounded in $(0, 1)$. Hence $F(n, x)$ itself is unbounded in $(0, 1)$.

If we wish to consider the full interval of convergence $(-\rho, \rho)$, we observe firstly that, as a particular case of (30) proved below, $|D^n f(x^2)/n!|^{1/n}$ is bounded in $(-1, 1)$, if $|f_n(x)/n!|^{1/n}$ is bounded in $(0, 1)$; and secondly, that if $\{a_r\}$ is a positive sequence and

$$f(x) = \sum_{r=0}^{\infty} a_r x^r,$$

then

$$\begin{aligned} D^n f(x^2) &= \sum_{2r=n}^{\infty} \frac{(2r)!}{(2r-n)!} a_r x^{2r} \\ &> \sum_{r=n}^{\infty} \frac{(2r)!}{(2r-n)!} a_r x^{2r}, \text{ by omission of positive terms,} \\ &> \sum_{r=n}^{\infty} \frac{r!}{(r-n)!} a_r x^{2r}. \end{aligned}$$

† p. 222.

But
$$f_n(x) = \sum_{r=n}^{\infty} \frac{r!}{(r-n)!} a_r x^r.$$

Thus, at $x = 1$, $\qquad D^n f(x^2)/n! > f_n(x)/n!,$

i.e. the first is unbounded, if the second is.

Then, under case (i), if $f(x)$ once more denotes the special series of the worked example, we have that $F(n, x)$ is bounded

for the function $f(-x^2)$, in the full interval of convergence $(-1, 1)$;

for the function $f(x)$, only in $(-1, \sigma)$, where $0 < \sigma < 1$;

for the function $f(x^2)$, only in $(-\sigma', \sigma)$, where $0 < \sigma, \sigma' < 1$.

Under case (ii) we make use of the series for $\log(1+x^2)$ and we have that $F(n, x)$ is bounded

for the function $\log(1+x^2)$, in the full interval of convergence $(-1, 1)$;

for the function $\log(1+x^2)+f(x)$, only in $(-1, \sigma)$, where $0 < \sigma < 1$;

for the function $\log(1+x^2)+f(x^2)$, only in $(-\sigma', \sigma)$, where $0 < \sigma, \sigma' < 1$.

4. The converse theorem

Conversely, we can prove that the necessary conditions (6), (12) are also sufficient to secure that, throughout some interval, $f(x)$ is expressible as the sum-function of a power-series.

In the first place we have already seen that this power-series, if it exists, must be a Taylor's series

$$\sum_{r=0}^{\infty} \frac{(x-a)^r}{r!} f_r(a) \tag{15}$$

of the function. Thus, as preliminary conditions, we require that every $f_r(a)$ exist, which is covered by (6), and that the series have an interval of convergence $(-\rho, \rho)$, i.e. that

$$|f_n(a)/n!|^{1/n}$$

be bounded, which is covered by (12). The region of validity of the series is also restricted by the condition of convergence $|x| < \rho$.

But, if the Taylor's series converges, it still remains to show that it converges to the desired sum-function $f(x)$, and for this we require the full condition (12), which, of course, includes (6) by presumption.

I shall first prove a lemma to the effect that (13) is a consequence of (12), i.e. in precise terms that

(16) *If, in some interval I,*
$$|f_n(x)/n!|^{1/n}$$
is bounded and if, for some a in I,
$$\rho_a \equiv \underline{\lim}|f_n(a)/n!|^{-1/n},$$
then, for every positive $\sigma_a < \rho_a$, every $n >$ some $N(a)$ independent of x, and every x in I_a the region common to the intervals I and $)a-\sigma_a, a+\sigma_a($, we have
$$|f_n(x)/n!| < \sigma_a(\sigma_a - |x-a|)^{-(n+1)}.$$

I establish this lemma by induction. Suppose that for every x in I and every n
$$|f_n(x)/n!|^{1/n} < \rho^{-1}. \tag{17}$$
For a given a in I, a given $\sigma_a < \rho_a$, every $n > N(a)$ and *some* x in I_a, suppose that
$$|f_n(x)/n!| < \sigma_a(\sigma_a - |x-a|)^{-(n+1)}. \tag{18}$$
Now by Lagrange's formula, chapter V (29),
$$f_n(x+y) = \sum_{r=0}^{p-1} y^r f_{n+r}(x)/r! + R_p,$$
where
$$R_p = y^p f_{n+p}(x+\theta y)/p!.$$
Choose y with the same sign as $x-a$ and such that $x+y$, and therefore also $x+\theta y$, lies in I. Then, by (17),
$$|R_p| < \frac{(n+p)!}{p!}|y/\rho|^p \rho^{-n}.$$
Hence, under the further restriction $|y| < \rho$, $R_p \to 0$ as $p \to \infty$, i.e., if $x+y$ lies in I and $|y| < \rho$, then
$$f_n(x+y) = \sum_{r=0}^{\infty} y^r f_{n+r}(x)/r!.$$
If now we introduce the hypothesis (18), we have, so long as $n > N(a)$,
$$|f_n(x+y)| < \sum_{r=0}^{\infty} \frac{(n+r)!}{r!}|y|^r \sigma_a(\sigma_a - |x-a|)^{-(n+r+1)},$$
i.e.
$$< n!\,\sigma_a(\sigma_a - |x-a|)^{-(n+1)}\left\{1 - \frac{|y|}{\sigma_a - |x-a|}\right\}^{-(n+1)},$$
if
$$|y| < \sigma_a - |x-a|.$$
Since y was chosen with the sign of $x-a$, we may write this in the form
$$|f_n(x+y)| < n!\,\sigma_a(\sigma_a - |x+y-a|)^{-(n+1)},$$
the restrictions on y being that
$$|x+y-a| < \sigma_a, \qquad x+y \text{ lies in } I, \qquad |y| < \rho.$$
The restriction on n is still that $n > N(a)$, independent of y. Hence the hypothesis (18), if true for some x in I_a, is also true for any $x+y$ in I_a, if $|y| < \rho$.

Now at a itself the hypothesis becomes

$$|f_n(a)/n!| < \sigma_a^{-n}.$$

But, since $$\sigma_a < \rho_a = \underline{\lim}|f_n(a)/n!|^{-1/n},$$

we have $$|f_n(a)/n!|^{-1/n} > \sigma_a, \quad \text{if } n > \text{some } N(a),$$

i.e. the hypothesis is true at a itself. Hence, by induction, we can extend it (in a finite number of steps) to any x in I_a and the lemma is therefore proved.

We can now establish conditions sufficient for the expansibility of $f(x)$, in the form:

(19) *If, in some interval I,*

$$|f_n(x)/n!|^{1/n}$$

is bounded and if, for some a in I,

$$0 < \sigma_a < \underline{\lim}|f_n(a)/n!|^{-1/n},$$

then throughout the common part of the intervals I and $(a-\sigma_a, a+\sigma_a)$ we can expand $f(x)$ in the Taylor's series

$$f(x) = \sum_{r=0}^{\infty}(x-a)^r f_r(a)/r!.$$

For we have $$f(x) = \sum_{r=0}^{n-1}(x-a)^r f_r(a)/r! + R_n$$

where $$R_n = (x-a)^n(1-\theta)^{n-1}f_n\{a+\theta(x-a)\}/(n-1)!,$$

using Cauchy's form of the remainder.

With the data of (19) we may choose σ_a' such that

$$0 < \sigma_a < \sigma_a' < \underline{\lim}|f_n(a)/n!|^{-1/n}.$$

Then, applying the lemma with σ_a' in place of σ_a, we have, if $n > N(a)$,

$$|R_n| < n|x-a|^n(1-\theta)^{n-1}\sigma_a'\{\sigma_a'-\theta|x-a|\}^{-(n+1)},$$

i.e. $$< \frac{n\sigma_a'|x-a|}{(\sigma_a'-\theta|x-a|)^2}\left\{\frac{|x-a|-\theta|x-a|}{\sigma_a'-\theta|x-a|}\right\}^{n-1}$$

$$< \frac{n\sigma_a\sigma_a'}{(\sigma_a'-\sigma_a)^2}\left\{\frac{|x-a|-\theta|x-a|}{\sigma_a'-\theta|x-a|}\right\}^{n-1}.$$

Since $$0 < |x-a|-\theta|x-a| < \sigma_a'-\theta|x-a|,$$

$R_n \to 0$ as $n \to \infty$, and so precisely

$$f(x) = \sum_{r=0}^{\infty}(x-a)^r f_r(a)/r!,$$

and the theorem is established.

5. Analytic functions

What we have proved in the preceding section is briefly that, within an interval in which $|f_n(x)/n!|^{1/n}$ is bounded, every Taylor's series of $f(x)$ whose point of origin lies in the interval converges to $f(x)$ in its open interval of convergence. If a particular Taylor's series converges at either end-point of convergence, we can extend the property to that point, for then $f(x)$ and the Taylor's series are both continuous at the point and the equality, holding in every neighbourhood of the point, extends to the point itself. Hence, more fully, we may say that

(20) *In an interval in which $|f_n(x)/n!|^{1/n}$ is bounded every Taylor's series of $f(x)$ converges to $f(x)$ whenever it converges at all.*

Suppose that we have actually

$$|f(x)/n!|^{1/n} < M$$

in the interval (a, b). The Taylor's series whose point of origin is the point ξ of this interval has the radius of convergence

$$\rho = \varliminf |f_n(\xi)/n!|^{-1/n},$$

i.e.

$$\rho \geqslant M^{-1}.$$

Thus every Taylor's series has an interval of convergence whose breadth is at least $2M^{-1}$. We may therefore divide up the interval (a, b) into a finite number (m say) of intervals (a, a_1), (a_1, a_2),..., (a_{m-1}, b), of breadth not exceeding $2M^{-1}$ such that throughout the interval (a_s, a_{s+1}) we can represent $f(x)$ by the Taylor's series with point of origin $\frac{1}{2}(a_s+a_{s+1})$. In other words, $f(x)$ is completely represented in the interval (a, b) by a finite number m of power-series.

We are reminded of the distinction between a function and its analytical representation, already emphasized in chapter I § 7. Here the function retains its individuality, although needing to be represented in different sections of its domain by different power-series.

A function which (i) is completely defined in an interval by a *finite number* of its Taylor's series,† and (ii) can be represented in the interval by *any* of its Taylor's series with points of origin in the interval is said to be *analytic* in the interval. It follows at once from (20) that

(21) *A sufficient condition that $f(x)$ be analytic in an interval is that $|f_n(x)/n!|^{1/n}$ be bounded in the interval.*

† It is to be understood that each of these Taylor's series, in defining the function, is to be allowed to range over its whole interval of convergence. Moreover, these intervals of convergence should overlap to avoid any difficulties at their end-points.

Conversely, if $f(x)$ be analytic in an interval, then, by virtue of the property (i) in its definition, it is the sum-function of a power-series in each of a finite number of (overlapping) sub-intervals, and so, by (12), $F(n,x)$ is bounded in each of these sub-intervals and consequently in the complete interval. In other words,

(22) *A necessary condition that $f(x)$ be analytic in an interval is that $|f_n(x)/n!|^{1/n}$ be bounded in the interval.*

The condition is therefore both necessary and sufficient, or, as we may shortly put it, '$f(x)$ analytic' is exactly equivalent to '$F(n,x)$ bounded'.

6. Prolongation of analytic functions

It will have been observed that the property (i) of the analytic function has been found sufficient to prove $F(n,x)$ bounded and therefore $f(x)$ analytic. In other words, (i) entails (ii) and the two properties assigned to the analytic function are thus consequential and not cumulative.

It is a notable corollary of this fact that an analytic function can, in general, be defined throughout the interval in which it is analytic by a representation in power-series extending over only part of that interval. Suppose that the function is defined to be analytic in (A, B) and that it is represented by a solitary power-series whose interval of convergence is (a,b) lying within (A, B). Form all the Taylor's series of the function with points of origin in this interval. In general, the aggregate of their intervals of convergence is an interval (a', b'), extending beyond (a,b), but, let us suppose, lying within (A, B). Since the function is defined to be analytic in (A, B), it is represented by any of these Taylor's series. We may now form the Taylor's series with points of origin in (a', b'). Between them they will, in general, extend the representation of the function to some wider interval (a'', b''), and so on. Since the function is analytic in (A, B), the radii of convergence of all its Taylor's series formed with points of origin in (A, B) are bounded below by some positive number. We are therefore certain of filling up the analytic interval (A, B) in a finite number of steps. As an example consider the function defined by the logarithmic series

$$f(x) \equiv x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \dots.$$

Its interval of convergence is $)-1, 1)$. The Taylor's series formed with the end-point $+1$ as point of origin is

$$\log 2 + \frac{x-1}{2} - \frac{1}{2}\left(\frac{x-1}{2}\right)^2 + \frac{1}{3}\left(\frac{x-1}{2}\right)^3 - \dots.$$

This series has the wider interval of convergence $)-1,3)$. The Taylor's series formed with the end-point 3 as point of origin is

$$2\log 2+\frac{x-3}{4}-\frac{1}{2}\left(\frac{x-3}{4}\right)^2+\frac{1}{3}\left(\frac{x-3}{4}\right)^3-\ldots$$

having the interval of convergence $)-1,7)$. After n such steps the analytical representation of the function will have been extended over the interval $)-1,2^n-1)$. Thus a finite number of steps suffices to extend the representation of the function to any finite interval which lies to the right of the point -1.

The representation, it is to be remarked, never extends *across* the point -1, as is to be expected, since the point is an infinity of the function. More generally, if a function is defined as analytic in a domain made up of p distinct intervals, we have to be given a representation of the function in some part at least of each of these intervals. It is enough, in actual fact, to know the representation of the function in the neighbourhood of some point of each interval or, as a bare sufficiency, merely to know the values of every derivative at a single point in each interval.

This process of extending the representation of an analytic function is known as 'analytic continuation' or 'analytic prolongation'. It is fundamental in the theory of the analytic function.

It is worth while here to bring into comparison the several ideas of 'function', 'continuous function', 'analytic function'. In the function without qualification it is of the essence of the definition that the value of the function at any one point does not prejudice the value at any other point.

In the continuous function to be given the value of the function in every neighbourhood of a point is to be given the value at the point itself, for this value is the limit of the neighbouring values. More generally to be given the values of the function at all *rational* points, say, of an interval is to be given the values of the function at *all* points of the interval. We may say briefly that 'a continuous function can be prolonged over an infinitely small distance'.

In the analytic function, as we have seen, to be given the values of the function throughout an interval is to be given its values throughout every interval of which the given interval forms part and in which the function remains analytic. In other words, if a function is analytic in every finite interval, to be given its values throughout any finite interval†

† Since an analytic function is necessarily continuous, it is sufficient to be given its values at all rational points, say, of such an interval.

is to be given its values at all finite points. We may say briefly that 'an analytic function can be prolonged over a finite distance'.

The idea of 'analytic function' belongs more especially to the Theory of Functions of a Complex Variable. Indeed, that theory is essentially a theory of analytic functions. The theory of real analytic functions is, in fact, common ground of the theory of real functions and the theory of complex functions, for, given the analytic function of a real variable x defined over some domain in the form

$$f(x) \equiv \sum^{\infty} a_n x^n,$$

we can at once define, over a suitable domain, a corresponding function of the complex variable $z = x+iy$ in the form

$$f(z) \equiv \sum^{\infty} a_n z^n.$$

This complex function reduces, for real values of z, to the corresponding real function.† The theory of real analytic functions is thus deducible as a special case of the theory of complex functions. For that reason we do not need to go very fully into the theory of the real analytic function. We shall give in the following sections only enough of the theory to make it reasonably self-supporting.

7. Functions non-analytic in some neighbourhood

The simplest cause of departure from the analytic condition in the neighbourhood of a point is the non-existence of derivatives at the point. Thus, if we write

$$f(x) \equiv x^{n+\theta} \qquad (0 < \theta < 1), \tag{23}$$

or

$$f(x) \equiv x^{n+1}\log x, \tag{24}$$

it is clear that, at the origin, $f(x)$ has no derivative of order greater than n.

Less simply, consider the function

$$f(x) \equiv \sum_{r=1}^{\infty} (x-r^{-1})^r \sin^2(\log|x-r^{-1}|), \tag{25}$$

where $|x| < 1$ and where, at each point $x = r^{-1}$, the corresponding rth term is omitted.

At the special point $x = n^{-1}$ the nth term is differentiable $n-1$ times, but no more. At every other point the series is differentiable term by term as often as we please. One such differentiation gives

$$\sum_{r=1}^{\infty} (x-r^{-1})^{r-1}\{r\sin^2(\log|x-r^{-1}|)+\sin(2\log|x-r^{-1}|)\};$$

† We may regard this, if we like, as an even more striking case of 'prolongation' of the real function, namely to a complex domain of definition.

a second such differentiation gives

$$\sum_{r=1}^{\infty}(x-r^{-1})^{r-2}\{r(r-1)\sin^2(\log|x-r^{-1}|)+(2r-1)\sin(2\log|x-r^{-1}|)+$$
$$+2\cos(2\log|x-r^{-1}|)\};$$

and so on. It is thus clear that the original series (25) and its successive differentiated series are comparable term by term with the geometric series

$$\sum_{r=1}^{\infty} x^r$$

and its differentiated series. Thus, in a region which is contained in the interval $)-1,1($ and which excludes the special points $x = r^{-1}$, every differentiated series is uniformly convergent and so defines the corresponding derivative of $f(x)$. We therefore have that

$f'(x)$ exists in $)-1,1($, but not at $x = 1$,

$f''(x)$ exists in $)-1,\frac{1}{2}($, but not at $x = \frac{1}{2}$,

and generally

$f_n(x)$ exists in $)-1,n^{-1}($, but not at $x = n^{-1}$.

There is thus no interval enclosing $x = 0$ throughout which every nth derivative of $f(x)$ exists. In other words, the function (25) is not analytic in the neighbourhood of $x = 0$.

We proceed now to examples in which every $f_n(x)$ exists at and near $x = 0$, but the analytic condition is not satisfied. Following Pringsheim,† write

$$f(x) = \sum_{r=0}^{\infty} \frac{1}{r!(1+a^r x)} \qquad (a > 1;\ x \geq 0). \tag{26}$$

Term-by-term differentiation gives formally

$$f_n(x)/n! = (-)^n \sum_{r=0}^{\infty} \frac{a^{nr}}{r!(1+a^r x)^{n+1}}, \tag{27}$$

which, since x is positive, is comparable with the convergent series of positive constants

$$\sum_{r=0}^{\infty} \frac{a^{nr}}{r!}.$$

Thus the formal equation (27) is valid and every derivative exists in the stated region $x \geq 0$. In particular

$$f_n(0)/n! = (-)^n \exp(a^n),$$

and so $\qquad \rho_0 = \lim|f_n(0)/n!|^{-1/n} = \lim \exp(-a^n/n) = 0.$

The Taylor's series with point of origin at $x = 0$ consequently exists, but converges only in a zero interval. The function cannot therefore

† *Math. Annalen*, 42 (1893), 153–84 (159). Other similar examples of non-analytic functions are given in this paper.

be analytic in the neighbourhood of $x = 0$. This is otherwise clear, for what we have just proved is in effect that $|f_n(0)/n!|^{1/n}$ is unbounded and therefore more generally that $|f_n(x)/n!|^{1/n}$ is unbounded in the neighbourhood of $x = 0$.

Lastly, consider the function

$$\left.\begin{aligned} f(x) &\equiv \exp(-x^{-2}) \qquad (x \neq 0) \\ f(0) &= 0 \end{aligned}\right\}. \tag{28}$$

Then
$$f'(x) = 2x^{-3}\exp(-x^{-2}),$$
$$f''(x) = (4x^{-6}-6x^{-4})\exp(-x^{-2}),$$
and generally
$$f_n(x) = (2x^{-3})^n\phi(x)\exp(-x^{-2}),$$
where $\phi(x)$ is a polynomial in x of degree not exceeding $2n-2$. Since $\exp(x^{-2})$ at the origin has an infinity of order higher than that of any x^n, it follows that, for every n,
$$f_n(0) = 0.$$
Thus Taylor's series with $x = 0$ as point of origin exists and converges everywhere, being everywhere zero. It thus represents the function $f(x)$ at $x = 0$ but nowhere else, and consequently $f(x)$ is not analytic in the neighbourhood of $x = 0$.

Actually, as $n \to \infty$, we find near $x = 0$ that $|f_n(x)/n!|^{1/n}$ is comparable with
$$2(x^{-3}/n)\exp(-x^{-2}/n),$$
which is unbounded near $x = 0$, as we can see, if we make $x \to 0, n \to \infty$ in such a way that nx^2 is constant.

8. Two lemmas on analytic functions

It is to be expected on psychological grounds that the elementary functions are generally analytic. For the restricted categories of function such as 'continuous', 'differentiable', 'analytic' have been introduced into analysis with the desire of extending characteristic properties of the elementary functions to as wide a class of functions as possible. Before discussing the analytic character of the elementary functions it is convenient to establish two lemmas on analytic functions. The first of these is:

(29) *If a function is analytic in an interval, its derived function is also analytic in the interval. Conversely, a function whose derived function is analytic in an interval is itself analytic in the interval.*

For, if $f(x)$ be analytic in some interval (a, b), then, in that interval and for some M,
$$|f_n(x)| < n!\,M^n.$$

Write in the first case $\qquad \phi(x) \equiv f'(x)$

and in the converse case $\qquad \psi'(x) \equiv f(x)$,

so that $\qquad |\phi_n(x)/n!|^{1/n} = |f_{n+1}(x)/n!|^{1/n} < [(n+1)M^{n+1}]^{1/n}$,

and $\qquad |\psi_n(x)/n!|^{1/n} = |f_{n-1}(x)/n!|^{1/n} < [M^{n-1}/(n-1)]^{1/n}$.

But $\qquad [(n+1)M^{n+1}]^{1/n}, \qquad [M^{n-1}/(n-1)]^{1/n}$

are both bounded, since, as $n \to \infty$, they both converge (to M). Thus

$$|\phi_n(x)/n!|^{1/n}, \qquad |\psi_n(x)/n!|^{1/n}$$

are both bounded in (a,b), i.e. $\phi(x)$, $\psi(x)$ are both analytic in (a,b).

The second lemma is briefly that 'an analytic function of an analytic function is itself an analytic function'. In precise terms we enunciate it thus:

(30) *If $y(x)$ is analytic in (a,b), if A, B are its extreme values in this interval, so that y lies in (A,B) when x lies in (a,b), and if $f(y)$ is an analytic function of y in (A,B), then $f\{y(x)\}$ is an analytic function of x in (a,b).*

If D denote differentiation in x, we have, by chapter V (38),

$$\frac{D^n f(y)}{n!} = \sum \frac{(Dy)^\alpha \left(\frac{D^2 y}{2!}\right)^\beta \cdots \left(\frac{D^n y}{n!}\right)^\theta}{\alpha!\, \beta! \ldots \theta!} f_r(y),$$

where $\alpha, \beta, \ldots, \theta, r$ are positive integers or zero satisfying the two equations

$$\left.\begin{aligned} \alpha+\beta+\ldots+\theta &= r \\ \alpha+2\beta+\ldots+n\theta &= n \end{aligned}\right\}. \tag{31}$$

Since $y(x)$ is analytic in (a,b), we have in this interval, for some M and every s,

$$|D^s y/s!| < M^s.$$

Since $f(y)$ is analytic in (A,B), we have, if x lies in (a,b), for some N and every r,

$$|f_r(y)| < r!\,N^r.$$

Hence $$\left|\frac{D^n f(y)}{n!}\right| < \sum \frac{r!}{\alpha!\,\beta! \ldots \theta!} M^{\alpha+2\beta+\ldots+n\theta} N^r,$$

i.e. $$< \sum \frac{r!}{\alpha!\,\beta! \ldots \theta!} M^n N^r.$$

Now fix r and effect the partial summation in $\sum$ by giving $\alpha, \beta, \ldots, \theta$ all values which satisfy (31). The partial sum is seen to be the coefficient of t^n in the expansion of the multinomial

$$(t+t^2+\ldots+t^n)^r,$$

i.e. in the expansion of $t^r(1-t)^{-r}$. This coefficient is

$$\frac{(n-1)!}{(r-1)!\,(n-r)!}.$$

For the complete summation $\sum$ we may therefore write

$$M^n \sum_{r=1}^{n} \frac{(n-1)!}{(r-1)!\,(n-r)!} N^r = M^n N(1+N)^{n-1}.$$

Since $M, N > 0$, we have, widening the inequality,

$$|D^n f(y)/n!| < M^n(1+N)^n,$$

i.e. $$|D^n f(y)/n!|^{1/n} < M(1+N) \quad \text{in } (a,b),$$

and the lemma is proved.

9. Analytic character of the elementary functions

Passing to the elementary functions, we have at once that

(32) *A polynomial is analytic in any finite interval.*

For, if the polynomial is of degree n, every derivative after the nth is zero and every derivative up to the nth is a polynomial and so bounded in any finite interval. Thus the set of all the functions $F(n,x)$ is also bounded. The proposition is otherwise obvious, since a polynomial can be expanded in a (terminating) series of powers of any $x-a$.

For the fractional and the irrational algebraic functions we consider the particular function
$$f(x) = x^p, \tag{33}$$
where p is not a positive integer. If p is negative, the function itself, and therefore also every derived function, has an infinity at $x = 0$. If p is positive, the derivatives of order exceeding p have all an infinity at $x = 0$. Thus in either case $f(x)$ is not analytic in the neighbourhood of the origin.

If we exclude the origin by writing, say,
$$0 < b < x < a, \tag{34}$$
we have
$$|f_n(x)/n!|^{1/n} = \left|\frac{(n-p-1)(n-p-2)...(-p)}{n(n-1)...1}\right|^{1/n} x^{p/n-1}.$$

Now, by the inequalities imposed on x, we have

$$\begin{aligned} x^{p/n-1} &< a^{p/n}b^{-1} \qquad (p > 0), \\ &< b^{p/n}b^{-1} \qquad (p < 0). \end{aligned}$$

As $n \to \infty$, each of these superior bounds converges (to b^{-1}). Hence $x^{p/n-1}$ is bounded in the stated interval (a, b).

Again, the coefficient of $x^{p/n-1}$ may be written in the form

$$(u_1 u_2 \ldots u_n)^{1/n},$$

where
$$u_n = \left|\frac{n-p-1}{n}\right| \to 1, \quad \text{as } n \to \infty.$$

Hence, by the theory of Cesàro's means,†

$$(u_1 u_2 \ldots u_n)^{1/n} \to 1, \quad \text{as } n \to \infty,$$

i.e. the coefficient is bounded. Therefore $F(n, x)$ itself is bounded in (a, b). Thus

(35) *If p is not a positive integer, the function x^p is not analytic in any neighbourhood of the origin, but is analytic in any closed interval which excludes the origin.*

More generally, in virtue of (30), we can say that

(36) *Any algebraic function expressible by means of surds and fractions is analytic in any closed interval which excludes the zeros of the surds and the infinities of the fractions.*

The elementary transcendental functions

$$e^x, \qquad \sin x, \qquad \cos x$$

are evidently analytic in any finite interval for their nth derivatives are respectively

$$e^x, \qquad \sin(x+\tfrac{1}{2}n\pi), \qquad \cos(x+\tfrac{1}{2}n\pi),$$

and therefore $F(n, x) \to 0$ as $n \to \infty$, whatever x. Alternatively, we may use the fact that they can all be expanded in powers of x in series which converge in any finite interval.

It follows, too, that the hyperbolic functions

$$\sinh x, \qquad \cosh x, \qquad \tanh x, \qquad \operatorname{sech} x$$

are analytic in any finite interval and that the hyperbolic functions

$$\coth x, \qquad \operatorname{cosech} x$$

and the trigonometrical functions

$$\tan x, \qquad \cot x, \qquad \sec x, \qquad \operatorname{cosec} x$$

are each analytic in a closed interval which excludes their infinities.

For the inverse functions

$$\log x, \qquad \sin^{-1} x, \qquad \tan^{-1} x, \qquad \sec^{-1} x$$

we consider the derivatives

$$\frac{1}{x}, \qquad \frac{1}{\sqrt{(1-x^2)}}, \qquad \frac{1}{1+x^2}, \qquad \frac{1}{x\sqrt{(x^2-1)}}.$$

† Cf. Bromwich, *Theory of Infinite Series* (1908), 383 (§ 154).

In virtue of (36) these four algebraic functions are analytic in any closed interval which excludes their infinities (if any). Hence, by (29), $\tan^{-1}x$ is analytic in any finite interval; $\log x$ is analytic in any closed interval which excludes $x = 0$; $\sin^{-1}x$ is analytic in any closed interval which excludes $x = \pm 1$; $\sec^{-1}x$ is analytic in any closed interval which excludes $x = 0, \pm 1$.

10. Zeros of analytic functions

In considering the function (28), namely,

$$f(x) \equiv \exp(-x^{-2}) \quad (x \neq 0), \qquad f(0) = 0,$$

which is non-analytic in the neighbourhood of the origin, we saw that the function itself and every derivative vanishes at the origin, although the function itself vanishes nowhere else than at the origin. This is in marked contrast to the analytic function for which we have the following characteristic property:

(37) *If a function is analytic in an interval and if at one (internal) point of the interval the function and every derivative vanish, then the function vanishes throughout the interval.*

For, if we form the Taylor's series with the point in question as origin, every coefficient in the series vanishes and therefore the series converges to zero for every x. The series represents the function throughout the interval, since the function is analytic throughout the interval; consequently the function vanishes throughout the interval.

It is instructive to compare such an analytic function with the type of function occurring in the dynamics of a particle moving rectilinearly in a field of force which is a function of position. If $f(t)$ denotes its displacement at time t, suppose that at some instant t_0 the velocity $f'(t)$ and the acceleration $f''(t)$ both vanish. Then the particle has come to rest and is under no acceleration; it therefore remains at rest in the position $f(t_0)$. In other words, $f(t) - f(t_0)$ vanishes identically provided only that its first *two* derivatives vanish at $t = t_0$. Evidently the class of such functions $f(t)$ is a more restricted class than the class of analytic functions.

From (37) we deduce an important corollary:

(38) *A function which is analytic in an interval and vanishes throughout a sub-interval, vanishes also throughout the complete analytic interval.*

For, if $f(x)$ vanish throughout (a, b), then every derivative also vanishes throughout (a, b), and therefore, by (37), $f(x)$ vanishes throughout the interval in which it is analytic.

It follows from this corollary that two analytic functions are identical, if they are identical in some sub-interval of the analytic interval, i.e. that an analytic function is uniquely determined by its values in any such sub-interval. This is, of course, no more than the principle, already employed, which forms the basis of analytic prolongation.

We also deduce from (37) an important proposition concerning the zeros of an analytic function:

(39) *The zeros of an analytic function are isolated.*

For suppose, on the contrary, that ξ is a non-isolated zero of $f(x)$ which is analytic in the neighbourhood of ξ. By this we mean that we can find zeros of $f(x)$ as close as we please to ξ and, in particular, we can find a sequence of zeros $\{x_n\}$ converging to ξ. Now

$$\frac{f(x)-f(\xi)}{x-\xi} \to f'(\xi), \quad \text{as } x \to \xi.$$

If we let x converge to ξ along the particular sequence $\{x_n\}$, the ratio on the left is always zero and hence the limit $f'(\xi)$ is zero.

Again, by Rolle's theorem, since $f(x)$ vanishes at ξ and at x_n, its derivative $f'(x)$ vanishes at an intermediate point x'_n. Thus $f'(x)$ has also a sequence of zeros $\{x'_n\}$ which converges to ξ, since the sequence $\{x_n\}$ converges to ξ. Hence, by the foregoing argument, $f''(\xi)$ vanishes. By successive application of this argument it follows inductively that every derivative of $f(x)$ vanishes at ξ and therefore, by (37), that $f(x)$ vanishes throughout its analytic interval. Accordingly, if we disregard this trivial case, we have the theorem stated.

11. Wronskians of analytic functions

It is as a consequence of these two theorems (37), (39) that analytic functions are especially favourable subjects for the characteristic processes of the Differential Calculus. Thus the methods of chapter VII for resolving an indeterminate form $f(x)/g(x)$ at a common zero ξ of $f(x)$, $g(x)$ fail, only if every derivative of $f(x)$, $g(x)$ vanishes at ξ. Accordingly, if f, g are analytic in the neighbourhood of the point in question, this can happen, only if f, g vanish throughout the interval, in which case the ratio f/g is devoid of meaning.

It follows, too, that, if $f(x)$, $g(x)$ are analytic near a common zero ξ, then $f'(x)$, $g'(x)$, being themselves analytic near ξ, by (29), cannot have a common zero in every neighbourhood of ξ, and therefore without exception

$$\lim_{x\to\xi}\frac{f(x)}{g(x)} = \lim_{x\to\xi}\frac{f'(x)}{g'(x)},$$

if the second limit exists.

Again, the theory connecting vanishing Wronskians and linear dependence developed in chapter V § 3, can be simplified, if the functions concerned are analytic. Precisely, we prove that

(40) *If $f_1(x),\ldots,f_n(x)$ are analytic in an interval and if their Wronskian*

$$(1,D)^{n-1}|f_1,\ldots,f_n|$$

vanishes throughout the interval, then the functions are linearly connected in the interval.

We have already proved in chapter V (19) that, if the Wronskian

$$(1,D)^{n-1}|f_1,\ldots,f_n|$$

vanishes throughout an interval, then $f_1,\ldots,f_n$ are connected by a linear relation

$$A_1 f_1+\ldots+A_n f_n = 0 \tag{41}$$

in any sub-interval in which one minor of the Wronskian, say

$$(1,D)^{n-2}|f_2,\ldots,f_n|, \tag{42}$$

does not vanish.

Suppose now that $f_1,\ldots,f_n$ are all analytic in the interval (a,b). Then, by (29), their derivatives are all analytic in the interval and, by (30), the minor (42) of the Wronskian is analytic. It thus either has isolated zeros ξ_1, $\xi_2\ldots$ in (a,b) or vanishes throughout (a,b).

In the former case $f_1,\ldots,f_n$ are connected by a linear relation, say by the relation (41), throughout the interval $)\xi_1,\xi_2($ and hence, by (38), throughout the analytic interval (a,b).

In the latter case we deal with the Wronskian (42) as we have just dealt with the original Wronskian. We can accordingly assert either that there is a linear relation

$$B_2 f_2+\ldots+B_n f_n = 0 \tag{43}$$

throughout (a,b), or else that throughout (a,b)

$$(1,D)^{n-3}|f_3,\ldots,f_n| = 0.$$

The linear relation (43) is, of course, a particular case of the linear relation (41) with A_1 zero. Proceeding inductively in this way we can therefore assert throughout (a,b) either the existence of a linear relation or, at length, the vanishing of the last Wronskian, i.e.

$$f_n = 0,$$

which is itself a linear relation. This proves the theorem. We see therefore that

(44) *With analytic functions the vanishing of the Wronskian throughout an interval is both necessary and sufficient for linear dependence in the interval.*

12. Analytic functions of many variables

If for $f(x)$, the function of one variable, we write the analytic condition in the form

$$|f_n(x)| < n!\,AB^n,$$

where A, B are independent of x, n, we are naturally led to consider for $f(x,y)$, the function of two variables, the analogous conditions

$$|f_{m,n}(x,y)| < m!\,n!\,AB^mC^n \tag{45}$$

and

$$|f_{m,n}(x,y)| < (m+n)!\,AD^{m+n}, \tag{46}$$

where $f_{m,n}(x,y)$ denotes the partial derivative $\partial^{m+n}f/\partial x^m\partial y^n$, and A, B, C, D are independent of x, y, m, n.

Now $(m+n)!/m!\,n!$ is a single term in the expansion of $(1+1)^{m+n}$, and we therefore certainly have that

$$1 \leqslant \frac{(m+n)!}{m!\,n!} < 2^{m+n}.$$

Then, if $B < C$, say,

$$m!\,n!\,AB^mC^n < (m+n)!\,AC^{m+n},$$

and

$$> (m+n)!\,A(\tfrac{1}{2}B)^{m+n}.$$

The inequalities (45), (46) are therefore equivalent and we may pass freely from one to the other. Of the two (46) is usually the simpler, and we may evidently state it in the form: '$|f_{m,n}(x,y)/(m+n)!|^{1/(m+n)}$, regarded as a function of x, y, m, n, is bounded.'

When, however, we attempt to construct a theory of analytic functions of two variables on the basis of these inequalities, we are confronted with the considerable difficulties of the convergence of double power-series $\sum\sum a_{m,n}x^my^n$. These difficulties centre primarily in the precise determination of the region of convergence, complicated by the fact that this region depends, in general, on the manner in which the terms are grouped in the summation. An adequate discussion of the convergence of double series would, of course, be out of place in this context, and we shall content ourselves with the statement of the restricted theorem:

(47) *If $f(x,y)$ denote the double power-series*

$$f(x,y) = \sum c_{m,n}x^my^n$$

which converges throughout the rectangle $|x| \leqslant a$, $|y| \leqslant b$, then throughout any interior rectangle $|x| \leqslant a' < a$, $|y| \leqslant b' < b$ every partial derivative $f_{m,n}(x,y)$ exists and we can write

$$|f_{m,n}(x,y)/m!\,n!| < AB^mC^n,$$

where A, B, C are independent of x, y, m, n.

We prove this theorem by arguments similar to those used in the proof of theorems (6), (12) of this chapter.

For the converse we can prove, in analogy with (16), (19) above, first the lemma:

(48) *If, in some interval I,*

$$|f_{m,n}(x,y)/(m+n)!|^{1/(m+n)}$$

is bounded, and if, for some (a,b) in I,

$$|f_{m,n}(a,b)/(m+n)!|^{1/(m+n)} \leqslant \rho_{a,b}$$

for every $m,n >$ some $M(a,b)$, then for every positive $\sigma_{a,b} < \rho_{a,b}$, every $m,n > M(a,b)$, and every x in the region common to I and the rectangle $|x-a|+|y-b| < \sigma_{a,b}$,

$$|f_{m,n}(x,y)/(m+n)!|^{1/(m+n)} < \sigma_{a,b}(\sigma_{a,b}-|x-a|-|y-b|)^{-(m+n+1)};$$

and then the main theorem:

(49) *If, in some interval I,*

$$|f_{m,n}(x,y)/(m+n)!|^{1/(m+n)}$$

is bounded, and if, for some (a,b) in I,

$$|f_{m,n}(a,b)/(m+n)!|^{1/(m+n)} < 1/\rho_{a,b}$$

for every $m,n >$ some $M(a,b)$, then we can expand $f(x,y)$ in the double Taylor's series†

$$f(x,y) = \sum_{r,s=0}^{\infty} (x-a)^r(y-b)^s f_{r,s}(a,b)/r!\,s!$$

throughout the region common to the interval I and the rectangle

$$|x-a|+|y-b| \leqslant \textit{any } \sigma_{a,b} < \rho_{a,b}.$$

These results can evidently be extended to functions of many variables.

WORKED EXAMPLE

If $\qquad f(x) = \sum_{r=0}^{\infty} u_r x^r,\;$ *where* $\;u_r = \dfrac{e^{-\sqrt{(r+1)}}}{r+1}+\dfrac{e^{-\sqrt{(r+2)}}}{r+2}+\ldots$ to ∞,

prove that

(i) *every $f_n(x)$ exists throughout the closed interval of convergence $(-1,1)$;*

(ii) *$F(n,x) = |f_n(x)/n!|^{1/n}$ is bounded in any interval $(-1,\sigma)$, where $0<\sigma<1$;*

(iii) *$F(n,x)$ is unbounded in the full interval of convergence $(-1,1)$.*

(i) Let m be the positive integer such that

$$m^2 < r+1 < (m+1)^2,$$

† From the form of Taylor's remainder it is evident that summation here is to be by 'diagonals', i.e. that we sum first for $r+s$ constant $(=N)$ and then for N (from 0 to ∞).

then, since $r^{-1}e^{-\sqrt{r}}$ decreases steadily as r increases, we have by the argument of Cauchy's 'condensation test'† that

$$u_r < (2m+1)\frac{e^{-m}}{m^2}+(2m+3)\frac{e^{-(m+1)}}{(m+1)^2}+\ldots \text{ to } \infty$$
$$< \frac{2m+1}{m^2}e^{-m}(1+e^{-1}+e^{-2}+\ldots \text{ to } \infty),$$

since $(2m+2p+1)/(m+p)^2 < (2m+1)/m^2$, if $p > 0$;

i.e.
$$u_r < \frac{e}{e-1}\left(\frac{2m+1}{m^2}e^{-m}\right).$$

By a similar argument

$$u_r > (2m+3)\frac{e^{-(m+2)}}{(m+2)^2}+(2m+5)\frac{e^{-(m+3)}}{(m+3)^2}+\ldots \text{ to } \infty$$
$$> (2m+3)\frac{e^{-(m+2)}}{(m+2)^2}.$$

Hence, as $r \to \infty$, u_r is of the order of $m^{-1}e^{-m}$, i.e. of

$$r^{-\frac{1}{2}}e^{-\sqrt{r}}.$$

Thus $|u_r|^{-1/r} \to 1$ as $r \to \infty$, and so the interval of convergence of $f(x)$, and therefore of every differentiated series $f_n(x)$, is the interval $(-1, 1)$.

Now
$$f_n(x) = \sum_{r=n}^{\infty} r(r-1)\ldots(r-n+1)u_r x^{r-n}.$$

This converges at $x = \pm 1$, since the series

$$\sum r(r-1)\ldots(r-n+1)u_r$$

is comparable, term by term, with such a convergent series as

$$\sum e^{-\frac{1}{2}\sqrt{r}}.$$

Thus every $f_n(x)$ exists throughout the closed interval of convergence $(-1, 1)$.

(ii) It is enough to consider the interval $(-1, 0)$, since, by (12) above, we already know that $F(n, x)$ is bounded in any interval $(0, \sigma)$, where $0 < \sigma < 1$. Write then $x = -\xi$ where $0 \leqslant \xi < 1$, and define

$$\phi_n(\xi) = \frac{f_n(-\xi)}{n!} = u_n-(n+1)\xi u_{n+1}+\frac{(n+1)(n+2)}{2!}\xi^2 u_{n+2}-\ldots \text{ to } \infty, \tag{1}$$

$$\psi_n(\xi) = \phi_0(\xi)-(n|1)(1+\xi)\phi_1(\xi)+(n|2)(1+\xi)^2\phi_2(\xi)-\ldots+(-)^n(1+\xi)^n\phi_n(\xi), \tag{2}$$

where the symbols $(n|s)$ denote binomial coefficients.

If E is the algebraic operator which changes u_r into u_{r+1}, we can write concisely

$$\phi_n(\xi) = (1+\xi E)^{-(n+1)}E^n u_0,$$
$$\psi_n(\xi) = (1+\xi E)^{-(n+1)}[(1+\xi E)-(1+\xi)E]^n u_0$$
$$= (1+\xi E)^{-(n+1)}(1-E)^n u_0.$$

This last equation has involved a rearrangement of terms in the infinite series for $\psi_n(\xi)$, but since no term has had to be shifted more than n places, questions of absolute convergence do not arise. Expanding the last expression for $\psi_n(\xi)$, we have

$$\psi_n(\xi) = (1-E)^n u_0-(n+1)\xi(1-E)^n u_1+\frac{(n+1)(n+2)}{2!}\xi^2(1-E)^n u_2-\ldots \text{ to } \infty. \tag{3}$$

We shall show that this is an alternating series of positive terms which steadily

† Cf. Bromwich, loc. cit., 23-4.

decrease to zero. The terms certainly tend to zero, since $\psi_n(\xi)$ is convergent, being made up of the sum of n convergent series. Hence, if the terms are steadily decreasing, they are certainly all positive. To prove them steadily decreasing it is sufficient, since $0 \leqslant \xi \leqslant 1$, to prove that, for every $r \geqslant 1$,

$$r(1-E)^n u_{r-1} > (n+r)(1-E)^n u_r.$$

Now

$$\begin{aligned} r(1-E)^n u_{r-1} - (n+r)(1-E)^n u_r \\ &= r[(1-E)^{n-1}u_{r-1} - (1-E)^{n-1}u_r] - (n+r)[(1-E)^{n-1}u_r - (1-E)^{n-1}u_{r+1}] \\ &= [r(1-E)^{n-1}u_{r-1} - (r+1)(1-E)^{n-1}u_r] - \\ &\qquad - [(n+r-1)(1-E)^{n-1}u_r - (n+r)(1-E)^{n-1}u_{r+1}] \\ &= (1-E)[r(1-E)^{n-1}u_{r-1} - (n+r-1)(1-E)^{n-1}u_r], \end{aligned}$$

if the influence of the operator E now extends to the *number* r as well as to the *suffix* r.

By successive use of this reduction we get at length

$$\begin{aligned} r(1-E)^n u_{r-1} - (n+r)(1-E)^n u_r &= (1-E)^n r(u_{r-1} - u_r) \\ &= (1-E)^n e^{-\sqrt{r}}, \end{aligned}$$

on substitution for u_{r-1}, u_r.

If $g(x)$ is any function possessing an nth derivative $g_n(x)$ and if $Eg(x) = g(x+1)$, then by repeated application of the formula of the mean we have

$$(1-E)^n g(x) = (-)^n g_n(x+\theta_1+\theta_2+\ldots+\theta_n), \tag{4}$$

where $\theta_1, \theta_2, \ldots, \theta_n$ all lie in $)0, 1($.

Now

$$De^{-\sqrt{x}} = -\tfrac{1}{2}x^{-\frac{1}{2}}e^{-\sqrt{x}},$$

$$D^2e^{-\sqrt{x}} = \tfrac{1}{4}(x^{-\frac{3}{2}} + x^{-1})e^{-\sqrt{x}},$$

and by induction we can prove that

$$D^n e^{-\sqrt{x}} = (-)^n(a_1 x^{-n+\frac{1}{2}} + a_2 x^{-n} + \ldots + a_n x^{-\frac{1}{2}n})e^{-\sqrt{x}},$$

where $a_1, a_2, \ldots, a_n$ are all positive numbers. Hence $D^n e^{-\sqrt{x}}$ has the sign $(-)^n$ and it follows from (4) that $(1-E)^n e^{-\sqrt{r}}$ is always positive. We have thus proved that the series (3) for $\psi_n(\xi)$ is an alternating series of positive terms steadily decreasing to zero. By the elementary theory of such series† we have

$$0 < \psi_n(\xi) < (1-E)^n u_0. \tag{5}$$

If now E_ϕ, E_ψ change ϕ_n, ψ_n into ϕ_{n+1}, ψ_{n+1} respectively, it follows from (2) that

$$E_\psi = 1-(1-\xi)E_\phi$$

and therefore that $\qquad (1+\xi)^n\phi_n(\xi) = (1-E_\psi)^n\psi_0(\xi),$

so that

$$\begin{aligned} (1+\xi)^n|\phi_n(\xi)| &< (1+E_\psi)^n\psi_0(\xi), \quad \text{since every } \psi \text{ is positive}, \\ &< (2-E)^n u_0, \quad \text{by (5)}, \\ &< (2+E)^n u_0, \quad \text{since } u_r \text{ is positive}, \\ &< 3^n u_0, \quad \text{since } u_r \text{ diminishes steadily}. \end{aligned}$$

Thus $\qquad |\phi_n(\xi)| < 3^n u_0, \quad \text{since} \quad 1+\xi \geqslant 1,$

i.e. $F(n, x)$ is bounded in the closed interval $(0, -1)$.

† Cf. Bromwich, loc. cit., 50–1 (§ 21).

(iii) To prove (iii) it is enough to show that $F(n, 1)$ is unbounded. Now

$$f_n(1)/n! = u_n+(n+1)u_{n+1}+\frac{(n+1)(n+2)}{2!}u_{n+2}+\dots \text{ to } \infty$$

$$= \frac{e^{-\sqrt{(n+1)}}}{n+1}+\{1+(n+1)\}\frac{e^{-\sqrt{(n+2)}}}{n+2}+\left\{1+(n+1)+\frac{(n+1)(n+2)}{2!}\right\}\frac{e^{-\sqrt{(n+3)}}}{n+3}+\dots$$

$$= \frac{1}{n+1}\left\{e^{-\sqrt{(n+1)}}+(n+1)e^{-\sqrt{(n+2)}}+\frac{(n+1)(n+2)}{2!}e^{-\sqrt{(n+3)}}+\dots\right\},$$

any rearrangement of terms being permissible, since all are positive. If we pick out those terms of the series with rational exponents, we have

$$f_n(1)/n! > \frac{1}{n+1}\sum_{r^2>n}^{\infty}\frac{(r^2-1)\dots(r^2-n)}{n!}e^{-r}$$

$$> \frac{1}{(n+1)!}\sum_{r=n+1}^{\infty}(r^2-1^2)(r^2-2^2)\dots\{r^2-(n-1)^2\}(r-n)re^{-r},$$

since $r^2-2 > r^2-2^2$, $\dots$, $r^2-(n-1) > r^2-(n-1)^2$, $r^2-n > (r-n)r$.

Thus
$$f_n(1)/n! > \frac{1}{(n+1)!}\sum_{r=0}^{\infty}\frac{(r+2n)!}{r!}e^{-(r+n+1)},$$

i.e.
$$> \frac{(2n)!}{(n+1)!}e^{-(n+1)}(1-e^{-1})^{-(2n+1)}$$
$$> n^{n-1}e^n(e-1)^{-2n}.$$

Hence
$$|f_n(1)/n!|^{1/n} > \frac{ne}{(e-1)^2}n^{-1/n},$$

which is unbounded, since $n^{-1/n} \to 1$ as $n \to \infty$.

EXAMPLES VIII

1. If, in some interval, $\quad |f_n(x)| < n!AB^n,$

show that in the same interval

$$|\delta^n f(x)| < n!AB|x|(1+B|x|)^{n-1} \qquad (\delta = xD),$$

and prove that, if $f(x)$ is analytic in a (finite) interval, it can be expanded in powers of e^x in that interval.

2. In a region in which $f(x)$ is analytic let $\rho(a)$ denote the radius of convergence of the Taylor's series expanded in powers of $x-a$, where a is a point of the region and $\rho(a)$ is supposed finite. Prove that $\rho(a)$ is a continuous function of a.

3. Show that the function

$$f(x) = \sum_{r=0}^{\infty}\frac{a^r}{r!(x^2+a^{2r})}$$

has derivatives of every order at every x, but that, if $a < 1$, the Taylor's series with point of origin $x = 0$ has a zero radius of convergence.

4. If, in some domain of $y(x)$, $z(x)$,

$$|D^r y/r!| < AM^r, \quad |D^r z/r!| < BN^r \qquad (0 < r),$$

show that
$$|D^n(yz)/n!| < AB\frac{M^{n+1}-N^{n+1}}{M-N},$$

and, if p is any positive integer, that

$$|D^n y^p/n!| < \frac{(n+p-1)!}{n!(p-1)!}A^pM^n \leq (n+1)^{p-1}A^pM^n.$$

Show also that
$$|D^n y^n/n!| < \tfrac{1}{2}(4AM)^n,$$

and, if $M < N$, that

$$|D^n y^p z/n!| < A^pBN^n(1-M/N)^{-p}.$$

5. If
$$u_r = D^r y/r! \qquad (r > 0),$$

and if $f(u_0, u_1, \ldots, u_n)$ is homogeneous of degree i and isobaric of weight w, show that $D^n f(u_0, u_1, \ldots, u_n)$ is homogeneous of degree i and isobaric of weight $w+n$. Show, moreover, that if c, C are respectively the sum of the coefficients and the sum of the *absolute values* of the coefficients of f, then the corresponding sums of coefficients in $D^n f$ are respectively

$$c\prod_{r=0}^{n-1}(w+i+r), \qquad C\prod_{r=0}^{n-1}(w+i+r).$$

6. If
$$(yD)^n z = \sum_{r=0}^{n-1} Y_{n,r} D^{n-r} z,$$

show that $Y_{n,r}$ written as a function of arguments $(D^r y/r!)$ is homogeneous of degree n and isobaric of weight r and that the sum of its coefficients $c_{n,r}$ satisfies the recurrence-formula
$$c_{n+1,r} = c_{n,r} + (n+r-1)c_{n,r-1}.$$

Prove that
$$c_{n,r} = \frac{(n+r-1)!}{(n-r-1)!\,r!\,2^r}.$$

7. If, in some interval of values of x,

$$|D^n y/n!| < AM^n, \qquad |D^n z/n!| < BN^n \qquad (n > 0),$$

show that, if $2N > M$,

$$|(yD)^n z/n!| < B\left(\frac{2AN^2}{2N-M}\right)^n$$

8. If p is a positive integer, show that

$$(y^pD)^n y$$

written as a function of arguments $(D^r y/r!)$ is homogeneous of degree $np+1$ and isobaric of weight n, and that the sum of its coefficients is

$$(p+2)(2p+3)\ldots\{(n-1)p+n\}.$$

If
$$(y^pD)^n z = \sum_{r=0}^{n-1} Y_{n,r} D^{n-r} z$$

and $Y_{n,r}$ is written as a function of arguments $(D^r y/r!)$, show that the sum of its coefficients $c_{n,r}$ satisfies the recurrence-formula

$$c_{n+1,r} = c_{n,r} + (np+r-1)c_{n,r-1}.$$

If $(y^pD)^n z$ be expressed as a function of arguments $(D^r y/r!)$, $(D^s z/s!)$, show that the sum of the coefficients is

$$(p+2)(2p+3)\ldots\{(n-1)p+n\}.$$

If, in some interval of values of x,

$$|D^n y/n!|,\ |D^n z/n!| < AM^n \qquad (n \geqslant 0),$$

and p is a positive integer, show that

$$|(y^p D)^n z/n!| < (p+1)^n A^{np+1} M^n.$$

9. Show that, if $y = e^x > 1$ or $y = x^a \geqslant 1$ ($a \gtrless 0$), then, as $n \to \infty$,

$$|(y^n D)^n y/n!|^{1/n} \to \infty.$$

10. If
$$D^n f(y) = \sum_{r=0}^{n-1} Y_{n,r} \frac{d^{n-r}}{dy^{n-r}} f(y),$$

show that $Y_{n,r}$, *written as a function of arguments* $(D^r y/r!)$, *is homogeneous of* degree $n-r$ and isobaric of weight n, and that the sum of its coefficients $c_{n,r}$ satisfies the recurrence-formula

$$c_{n+1,r} = c_{n,r} + (2n-r+1)c_{n,r-1}.$$

Prove that
$$c_{n,r} = \frac{n!(n-1)!}{(n-r)!(n-r-1)!r!}.$$

11. If, in some interval of values of x,

$$|D^n y/n!| < AM^n \qquad (n > 0)$$

and if also for the corresponding values of y

$$\left|\frac{1}{n!}\frac{d^n}{dy^n} f(y)\right| < BN^n \qquad (n \geqslant 0),$$

show that
$$|D^n f(y)/n!| < \frac{ABN}{AN+1}\{M(AN+1)\}^n.$$

Deduce an alternative proof of § 8 (30).

12. If
$$D^n \frac{1}{y} = \sum_{r=1}^{n} \frac{(-)^r Y_{n,r}}{y^{r+1}},$$

show that $Y_{n,r}$ written as a function of arguments $(D^r y/r!)$ is homogeneous of degree r and isobaric of weight n, and that the sum of its coefficients $c_{n,r}$ satisfies the recurrence formula
$$c_{n+1,r} = (n-r)c_{n,r} + rc_{n,r-1}.$$

Prove that
$$c_{n,r} = \frac{n!(n-1)!}{(r-1)!(n-r)!}.$$

13. Show that, if
$$\frac{y^{n+1}}{n!} D^n \frac{1}{y}$$

be written as a function of arguments $(D^r y/r!)$, the sum of its coefficients is zero and the sum of the *absolute values* of its coefficients is 2^{n-1}.

If, in some interval of values of x,

$$|D^n y/n!| < AM^n \quad (n > 0), \qquad |y| > B,$$

prove that
$$\left|\frac{1}{n!} D^n \left(\frac{1}{y}\right)\right| < \frac{A}{(A+B)B}\left\{\frac{M(A+B)}{B}\right\}^n.$$

14. If, in some closed interval of values of x, the function $y(x)$ is analytic and does not vanish, show that the reciprocal function $\{y(x)\}^{-1}$ is also analytic in the interval.

Show also how to deduce this result from § 8 (30), (35) above.

15. If $$\frac{d^nz}{dy^n} = \left(\frac{dx}{dy}\right)^n D^n z + \sum_{r=1}^{n-1}\sum_{s=1}^{r}(-)^s\left(\frac{dx}{dy}\right)^{n+s} Y_{n,r,s}D^{n-r}z,$$

show that $Y_{n,r,s}$ written as a function of arguments $(D^r y/r!)$ is homogeneous of degree s and isobaric of weight $r+s$ and that the sum of its coefficients $c_{n,r,s}$ satisfies the recurrence-formulae

$$c_{n+1,r,s} = c_{n,r,s}+(r+2s-1)c_{n,r-1,s}+2(n+s-1)c_{n,r-1,s-1} \qquad (r > 1),$$

$$c_{n+1,1,1} = c_{n,1,1}+2n.$$

Prove that $$c_{n,r,s} = \frac{(n+s-1)!(r-1)!}{(n-r-1)!(r-s)!s!(s-1)!}.$$

16. If $$\frac{d^nz}{dy^n} = \left(\frac{dx}{dy}\right)^n D^n z + \sum_{r=1}^{n-1}\left(\frac{dx}{dy}\right)^{n+r} Y_{n,r}D^{n-r}z,$$

show that $Y_{n,r}$ written as a function of arguments $(D^r y/r!)$ is homogeneous of degree r and isobaric of weight $2r$, and that the sum of its coefficients is

$$(-)^r\frac{n!(n-1)!}{(n-r)!(n-r-1)!r!}.$$

Show that the sum of the *absolute values* of its coefficients $C_{n,r}$ is the coefficient of t^{n-1} in

$$\frac{(n-1)!}{(n-r-1)!}(1+t)^n(2+t)^{r-1}.$$

Show *a fortiori* that, if t is any positive number,

$$C_{n,r} < \frac{(n-1)!}{(n-r-1)!}(1+t)^n t^{1-n}(2+t)^{r-1}.$$

17. If $$\frac{d^nx}{dy^n} = \sum_{r=1}^{n-1}(-)^r Y_{n,r}\left(\frac{dx}{dy}\right)^{n+r},$$

show that $Y_{n,r}$ written as a function of arguments $(D^r y/r!)$ is homogeneous of degree r and isobaric of weight $n+r-1$ and that the sum of its coefficients $c_{n,r}$ satisfies the recurrence-formula

$$c_{n+1,r} = (n+2r-1)c_{n,r}+2(n+r-1)c_{n,r-1}.$$

Show that, if $$\frac{1}{n!}\left(\frac{dy}{dx}\right)^{2n-1}\frac{d^nx}{dy^n}$$

be written as a function of arguments $(D^r y/r!)$, the sum of its coefficients is 1 and the sum of the absolute values of its coefficients is less than 6^n.

18. If, for values of x in the closed interval (a, b), the function $y(x)$ is analytic and the derivative $y'(x)$ does not vanish, prove that the corresponding inverse function $x(y)$ is analytic in the corresponding interval $\{y(a), y(b)\}$.

19. If, in some interval of values of x,

$$|D^n f(x)/n!| < AM^n, \qquad |D^n\phi(x)/n!| < BN^n \qquad (n > 0)$$

and $$|1-a\phi'(x)| > C,$$

show that we can find constants A_1, M_1 such that in the same interval

$$\left|\frac{1}{n!}\left\{\frac{1}{1-a\phi'(x)}\frac{d}{dx}\right\}^n\frac{f'(x)[\phi(x)]^n}{1-a\phi'(x)}\right| < A_1 M_1^n.$$

20. *If a continuous function $y(x)$ is defined implicitly by the relation*

$$y = a + x\phi(y)$$

and if, corresponding to values of x in some closed interval, $1 - x\phi'(y)$ does not vanish and $\phi(y), f(y)$ are analytic functions of y, show that in the same interval $f(y)$ and $\{1 - x\phi'(y)\}^{-1}$ are also analytic functions of x.

21. If in Lagrange's series

$$|D_a^n \phi(a)/n!| < AM^n,$$

show that the series for y certainly converges, if

$$|x| < 4AM,$$

and that, if also $\quad |D_a^n f(a)/n!| < BN^n \qquad (N > M),$

then the series for $f(y)$ certainly converges, if

$$|x| < \frac{AN^2}{N-M}.$$

22. Determine the number of real roots of the equation

$$ax^m - x + 1 = 0 \qquad (m > 2),$$

and show that, if $\quad a < \dfrac{(m-1)^{m-1}}{m^m},$

the equation has always a real root x_1 which converges to 1 as a converges to zero, and only one such root.

Show also that, if $\quad |a| <$ some $\eta < \dfrac{(m-1)^{m-1}}{m^m},$

this root x_1 is equal to the sum of the convergent series

$$\sum_{n=0}^{\infty} \frac{(mn)!}{n!(mn-n+1)!} a^n.$$

Show further that, if p be any positive integer, we have, under the same conditions,

$$(x_1)^p/p = \sum_{n=0}^{\infty} \frac{(mn+p-1)!}{n!(mn-n+p)!} a^n, \qquad \log x_1 = \sum_{n=1}^{\infty} \frac{(mn-1)!}{n!(mn-n-1)!} a^n.$$

23. If x, y are connected by the implicit relation $f(x, y) = 0$ and if we write

$$\frac{\partial^{p+q} f}{\partial x^p \partial y^q} = p!q! u_{p,q}, \qquad \frac{d^n y}{dx^n} = \sum_{r=1}^{2n-1} (-)^r Y_{n,r} \left(\frac{\partial f}{\partial y}\right)^{-r},$$

show that $Y_{n,r}$ written as a function of arguments $(u_{p,q})$ is homogeneous of degree r and isobaric of weight n in the suffixes p and of weight $r-1$ in the suffixes q.

Show also that the sum of its coefficients $c_{n,r}$ satisfies the recurrence-formula

$$c_{n+1,r} = (n+r)c_{n,r} + (3r-4)c_{n,r-1} + 2(r-2)c_{n,r-2}$$

and that C_n, the sum of the *absolute values* of the coefficients in

$$\left(\frac{\partial f}{\partial y}\right)^{2n-1} \frac{d^n y}{dx^n},$$

satisfies the recurrence-inequality

$$C_{n+1} < (13n-7)C_n.$$

24. If, in some interval of values of x, the relation $f(x, y) = 0$ defines an implicit function $y(x)$ such that $f(x, y)$ is 'analytic', i.e. such that for every $p, q \geqslant 0$ $\dfrac{\partial^{p+q}f}{\partial x^p \partial y^q}$ exists and

$$\left| \frac{\partial^{p+q}f}{\partial x^p \partial y^q} \right| < Ap!q!\, M^p N^q,$$

where A, M, N are numbers independent of x, y, p, q, show that numbers A_1, M_1 independent of n can be found such that

$$\left| \frac{1}{n!} \left(\frac{\partial f}{\partial y}\right)^{2n-1} \frac{d^n y}{\partial x^n} \right| < A_1 M_1^n.$$

25. If $f(x, y)$ is 'analytic' in some closed interval in which $f(x, y) = 0$ defines an implicit function $y(x)$ and if $f_y(x, y)$ does not vanish in this interval, show that $y(x)$ is analytic for the corresponding values of x.

IX

MAXIMA AND MINIMA

1. Recapitulation

It has been seen in chapter III that every bounded set, and, in particular, the set of values in an interval of a function bounded in the interval, has an upper bound and a lower bound. We have seen further that, if a function is continuous in an interval, it is necessarily bounded there and its bounds are actual values of the function at points of the interval. These bounds are therefore now the greatest and the least values of the function.

In chapter IV we have further seen that, if a function of one variable is differentiable in an interval (and therefore certainly continuous in the interval), the greatest and the least values can occur only at the end-points of the interval or at zeros of the derivative. On the other hand, we have also seen that zeros of the derivative are not necessarily points of greatest or least value.

In addition, we have defined a maximum (minimum) of a function as a point at which the function is at least as great (small) as elsewhere in the neighbourhood of the point (the 'neighbourhood' meaning, as always, an interval *completely enclosing* the point). We have shown that, if a function of one variable is differentiable in the neighbourhood of a maximum or a minimum, its derivative vanishes there. but conversely that the derivative may vanish at points neither maxima nor minima.

In the light of these facts it is well to distinguish clearly between three classes of points at which the values of a function of a single variable, differentiable in an interval, become notable among its values in the interval:

(i) the points of greatest and of least value in the interval (the points of *bounding* value);

(ii) the maxima and minima, if any (the points of *turning* value);

(iii) the zeros of the derivative, if any (the points of *stationary* value).

We should observe that, whereas bounding values must be defined with reference to a *given* interval, turning values and stationary values depend only on the behaviour of the function in an *arbitrary* neighbourhood of the point: the former are jointly functions of the function and a prescribed interval, the latter intrinsic functions of the function itself.

2. Turning values of the function of one variable

The points of bounding value of the function in the given interval must, from the definition, be also turning-points of the function or else end-points of the interval. The turning-points of the function are also points of bounding value in a properly chosen interval.

For a differentiable function turning values must be stationary values, but not all stationary values are turning values. The theory for functions of a single variable is completed by the following theorem that enables us, in general, to discriminate which stationary values are also turning values and therefore to determine the bounding values of the function in a prescribed interval:

(1) *The function being supposed sufficiently differentiable, a stationary value is also a turning value, if and only if the earliest derivative which does not vanish at the point is of even order. The turning value is a maximum or a minimum according as the value of this derivative at the point is negative or positive.*

To prove $f(a)$ a turning value of $f(x)$ we have to show that, for sufficiently small values of h,
$$f(a+h)-f(a)$$
has a sign independent of the sign of h. If $f_n(x)$ exists at $x = a$, then by Taylor's limit, chapter V (34),
$$\lim_{h\to 0}\left\{f(a+h)-\sum_{r=0}^{n-1}\frac{h^r}{r!}f_r(a)\right\}\Big/h^n = f_n(a)/n!.$$
Suppose now that
$$f_1(a) = 0, \quad ..., \quad f_{n-1}(a) = 0, \qquad f_n(a) \neq 0.$$
The limit reduces to
$$\lim_{h\to 0}\{f(a+h)-f(a)\}/h^n = f_n(a)/n!,$$
and so, if h is sufficiently small, $f(a+h)-f(a)$ has the sign of $h^n f_n(a)$. Hence, if n be even, $f(a+h)-f(a)$ has the sign of $f_n(a)$, which is independent of the sign of h and thus a is a turning-point of $f(x)$. Moreover, the turning value is a maximum or minimum according as this sign of $f(a+h)-f(a)$ is negative or positive, i.e. according as $f_n(a)$ is negative or positive.

On the other hand, if n be odd, $f(a+h)-f(a)$ has the sign of h and the function is neither maximum nor minimum. If $y = f(x)$ can be represented by a curve, this curve crosses its tangent at such a point, which is therefore an inflexion of the curve. We may borrow from geometry and conveniently term such a point an *inflexional* point.

Since the proof of (1) turns essentially on the existence of Taylor's

limit, we can evidently replace the 'derivative' of the enunciation by the 'umbral derivative' $[f_n(a)]$ of chapter V § 8, which is defined inductively by the recurrence-formula

$$[f_n(a)]/n! = \lim_{h\to 0}\left\{f(a+h)-\sum_{r=0}^{n-1}\frac{h^r}{r!}[f_r(a)]\right\}\Big/h^n.$$

The existence of the umbral derivative of any order, as we have seen in chapter V, imposes less stringent conditions than those imposed by the existence of the full derivative of equal order. We thus widen the applicability of (1), if we re-enunciate it in the form:

(2) *Granted the existence of sufficient umbral derivatives at the point, a stationary value is also a turning value, if and only if the earliest umbral derivative which does not vanish at the point is of even order. The turning value is a maximum or a minimum according as the value of this umbral derivative at the point is negative or positive.*

The tests of theorem (1) are applicable throughout any interval in which $f(x)$ is analytic, for then at every point of the interval every derivative exists and not every derivative vanishes—we exclude the trivial case in which $f(x)$ is constant in the interval. There is thus an earliest derivative not vanishing at the point, and the theorem provides an exact and complete solution to the problem of determining the turning values of the function in the interval.

In an interval in which the function is non-analytic the tests of both theorems (1) and (2) may be inapplicable. Thus the function

$$\left.\begin{aligned} f(x) &= \exp(-x^{-2}) \qquad (x\neq 0)\\ f(0) &= 0 \end{aligned}\right\} \tag{3}$$

has a minimum at the origin, since elsewhere it is positive. But the tests of theorems (1), (2) fail to indicate this minimum, since every derivative vanishes at the origin.

Again the function

$$\left.\begin{aligned} f(x) &= x^2\sin^2(x^{-1}) \qquad (x\neq 0)\\ f(0) &= 0 \end{aligned}\right\} \tag{4}$$

has a minimum at the origin, since $f(x)$ is everywhere positive or zero. But the tests of theorems (1), (2) again fail, since $f'(0) = 0$ and neither $f''(0)$ nor even $[f''(0)]$ exists.

3. Functions of many variables

When we pass on to functions of n variables, we base our discussion on the same initial principles. The 'interval' is now the n-dimensional interval

$$|x_1-a_1| < \delta_1,\dots, |x_n-a_n| < \delta_n.$$

If $f(x_1,\dots,x_n)$ be a function continuous throughout such an interval, its values therein will include a greatest value and a least value; these are the *bounding* values of the function in the interval. These values, unless they occur at points on the boundary of the interval, will be respectively a maximum and a minimum of the function. More generally, we define a point $(a_1,\dots,a_n)$ to be a minimum (maximum) of the function, if the point can be completely enclosed by some interval

$$|x_1-a_1|,\dots,|x_n-a_n| < \delta$$

throughout which

$$f(x_1,\dots,x_n)-f(a_1,\dots,a_n) \geqslant 0 \quad (\leqslant 0),$$

i.e. if, for all sufficiently small values of $|h_1|,\dots,|h_n|$,

$$f(a_1+h_1,\dots,a_n+h_n)-f(a_1,\dots,a_n)$$

is positive (negative) or zero, independently of the sign of $h_1,\dots,h_n$. These points $(a_1,\dots,a_n)$ are the *turning-points* of the function.

If all the first derivatives $f_1,\dots,f_n$ of f exist throughout the interval, we define a *stationary* point of the function to be a point at which *all* these first derivatives vanish, i.e. $(a_1,\dots,a_n)$ is a stationary point, if

$$f_1(a_1,\dots,a_n) = 0, \quad \dots, \quad f_n(a_1,\dots,a_n) = 0.$$

Since the maxima of $f(x_1,\dots,x_n)$ are precisely the minima of $-f(x_1,\dots,x_n)$, it will be sufficient in future discussion to consider minima alone. The corresponding results for maxima are at once deducible by changing the sign of f.

Suppose now that $(a_1,\dots,a_n)$ is a minimum of $f(x_1,\dots,x_n)$ and let us define the arguments $x_1,\dots,x_n$ of f as themselves functions of a single parameter t,

$$x_1 = x_1(t), \quad \dots, \quad x_n = x_n(t) \tag{5}$$

such that $(a_1,\dots,a_n)$ itself is included among the values of $(x_1,\dots,x_n)$, say for $t = t_0$. If, further, we make the arbitrary functions (5) continuous near t_0, then an interval

$$|t-t_0| < \epsilon \tag{6}$$

defines an interval

$$|x-x_1|,\dots,|x-x_n| < \delta \tag{7}$$

in the field of n variables.

The values of the function

$$F(t) \equiv f\{x_1(t),\dots,x_n(t)\}$$

in the interval (6) are included among the values of the function $f(x_1,\dots,x_n)$ in the interval (7). But, by definition, $f(a_1,\dots,a_n)$, i.e. $F(t_0)$, is the least of the latter set of values for proper choice of δ. Hence too $F(t_0)$ is the least of the set of values of $F(t)$ in (6) for proper choice of ϵ. In other words, t_0 is a minimum of $F(t)$.

We may call (5) a 'continuous curve' (in the field of n variables) passing through $(a_1,\ldots,a_n)$, and we may then say briefly that

(8) *A point which is a minimum of $f(x_1,\ldots,x_n)$ is necessarily also a minimum of f taken along any continuous curve through the point.*

We at once deduce that

(9) *The minima of a differentiable function are stationary points of the function.*

For, if we restrict the continuous curves (5) to be differentiable curves, i.e. to be such that the first derivatives $x_1'(t),\ldots,x_n'(t)$ all exist in the interval under discussion, then, by chapter VI (18),

$$F'(t) = x_1'(t)f_1(x_1,\ldots,x_n)+\ldots+x_n'(t)f_n(x_1,\ldots,x_n). \tag{10}$$

If $(a_1,\ldots,a_n)$, corresponding to t_0 on the curve, is a minimum of f, then t_0 is a minimum of F and therefore also a stationary point of F. Thus, by (10),

$$x_1'(t_0)f_1(a_1,\ldots,a_n)+\ldots+x_n'(t_0)f_n(a_1,\ldots a_n) = 0 \tag{11}$$

for every differentiable curve through the minimum. Choose in particular the curve

$$\left.\begin{aligned} x_r &= a_r \qquad (r = 1,\ldots,s-1,s+1,\ldots,n) \\ x_s &= t \end{aligned}\right\}, \tag{12}$$

where $t_0 = a_s$. This passes through $(a_1,\ldots,a_n)$ and is differentiable: in fact we have

$$x_r' = 0 \qquad (r = 1,\ldots,s-1,s+1,\ldots,n),$$

$$x_s = 1.$$

Hence, from (11), $\qquad f_s(a_1,\ldots,a_n) = 0.$

Taking in succession the n curves corresponding to $s = 1,\ldots,n$, we obtain as necessary conditions for the minimum

$$f_s(a_1,\ldots,a_n) = 0 \qquad (s = 1,\ldots,n),$$

which establishes the theorem.

Since the function of n variables includes the function of one variable as a particular case, it follows, as for the function of one variable, that not every stationary value is a turning value.

4. Conditions sufficient for a minimum

The principle of (8), as we have just seen in a particular case, enables us to deduce *necessary* conditions for minima of functions of many variables from the corresponding theory for functions of one variable.

We can in similar fashion deduce *sufficient* conditions from the same theory by relying on the converse of (8), namely, that

(13) *A point is a minimum of $f(x_1,..., x_n)$, if it is a minimum of f taken along* EVERY *continuous curve through the point.*

We shall in point of fact need the converse in the more stringent form in which 'every continuous curve' is replaced by a select class of such curves, namely those for which, in the defining equations (5), every umbral derivative $[D^p x_s(t)]$ of order p exists at the point, p being some specified integer, and, in addition, the first derivatives $x_s'(t)$ do not all vanish at the point.†

We may state the stricter converse in the form:

(14) *A point O is necessarily a minimum of $f(x_1,..., x_n)$, if, for some specified integer p, it is a minimum of f taken along every continuous curve*

$$x_s = x_s(t) \qquad (s = 1,..., n)$$

through O such that the umbral derivatives $[D^p x_s(t)]$ of order p all exist at O and the first derivatives $Dx_s(t)$ do not all vanish at O.

To prove this theorem we rely on the following lemma:

(15) *Given a point O and a sequence of points converging to O in a field of n variables and given a positive integer p, then we can choose a subsequence of the given points through which we can draw a continuous curve*

$$x_s = x_s(t) \qquad (s = 1,..., n)$$

passing also through O and such that the umbral derivatives $[D^p x_s(t)]$ of order p all exist at O and the first derivatives $Dx_s(t)$ do not all vanish at O.

It will be less wearisome to postpone the proof of the lemma and to show first how we make use of it in proving (14). Let us imagine therefore that the lemma is proved. Now, if $(\xi_1,..., \xi_n)$ be a minimum of $f(x_1,..., x_n)$, we can find some ϵ such that

$$f(x_1,..., x_n) \geqslant f(\xi_1,..., \xi_n),$$

if $$|x_1-\xi_1|,..., |x_n-\xi_n| < \epsilon.$$

Conversely therefore, if $(\xi_1,..., \xi_n)$, or P say, be not a minimum of f, then given *any* positive ϵ however small we can find *some* $x_1(\epsilon),..., x_n(\epsilon)$ such that $|x_1-\xi_1|,..., |x_n-\xi_n| < \epsilon$, but

$$f(x_1,..., x_n) < f(\xi_1,..., \xi_n). \tag{16}$$

If then we choose a positive sequence $\{\epsilon_N\}$ converging to zero, the

† We may say that such a curve is 'p times (umbrally) differentiable' at the point and has a determinate 'tangential direction' at the point. Evidently the satisfaction of the latter condition depends on the parameter defining the curve as well as on the intrinsic character of the curve itself.

corresponding sequence of points $\{x_1(\epsilon_N),..., x_n(\epsilon_N)\}$ converges to P and at every point of the sequence the inequality (16) is satisfied. By the lemma we can describe through these points and P a continuous curve having the stipulated derivatives at P. The inequality (16) is then satisfied at points of this curve arbitrarily near P, and therefore P is not a minimum of the values of f considered along this curve. Hence, if P is a minimum of f along *every* such curve through P, it must also be a minimum of the unrestricted function, i.e. of the function regarded n-dimensionally. Thus (14) is established, if the lemma holds.

5. Proof of the lemma

For brevity consider only the case $n = 2$, $p = 2$. Take the given point O as origin of polar coordinates (r,θ) in the plane of the points. The inequalities

$$\alpha \leqslant \theta \leqslant \beta \qquad (0 \leqslant \alpha < \beta \leqslant 2\pi)$$

define an infinite sector in the plane which we may call 'the sector (α,β)'.

Now it may happen that, however small α, the sector $(0,\alpha)$ contains always an infinity of points of the sequence. In other words, for every positive ϵ the sector $(-\epsilon, +\epsilon)$ contains always an infinity of points of the sequence.

If not, then for some α the sector $(0,\alpha)$ contains at most a finite number† of points of the sequence. Since the sector $(0,2\pi)$ contains all the points of the sequence, the set of numbers $\{\alpha\}$ is bounded and so has either a greatest number α_0 or an upper bound α_0. In either case α_0 is such that for every positive ϵ the sector $(0,\alpha_0-\epsilon)$ contains at most a finite number of points of the sequence, while the sector $(0,\alpha_0+\epsilon)$ contains an infinity of points of the sequence. In consequence the sector $(\alpha_0-\epsilon,\alpha_0+\epsilon)$ itself contains an infinity of points of the sequence. Thus in every case we can choose α_0 such that for every positive ϵ the sector $(\alpha_0-\epsilon,\alpha_0+\epsilon)$ contains an infinity of points of the sequence.

Now choose a positive sequence $\{\epsilon_n\}$ converging steadily to zero. Then from the given sequence of points we can pick out a subsequence whose polar coordinates (r_n, θ_n) satisfy the successive inequalities:

$$r_1 > r_2 > ... > r_n > ...$$

$$\alpha_0-\epsilon_1 < \theta_1 < \alpha_0+\epsilon_1, \quad ..., \quad \alpha_0-\epsilon_n < \theta_n < \alpha_0+\epsilon_n, \quad ....$$

This is always possible, since every sector $(\alpha_0-\epsilon_n,\alpha_0+\epsilon_n)$ contains an infinity of the given points and, in the given sequence, $r_n \to 0$ as $n \to 0$. Evidently in the chosen subsequence $\{(r_n,\theta_n)\}$ we shall have, as $n \to \infty$,

$$r_n \to 0 \text{ monotonically}, \qquad \theta_n \to \alpha_0.$$

† 'At most a finite number', i.e. a finite number or none at all.

Define the curve $x = x(r)$, $y = y(r)$, where, for $r_n \geqslant r \geqslant r_{n+1}$,

$$x(r) \equiv \frac{x_n(r-r_{n+1})+x_{n+1}(r_n-r)}{r_n-r_{n+1}},$$

$$y(r) \equiv \frac{y_n(r-r_{n+1})+y_{n+1}(r_n-r)}{r_n-r_{n+1}},$$

x_n, y_n being the cartesian coordinates of (r_n, θ_n). The curve is evidently a broken polygon of infinitely many sides with vertices at the points (x_n, y_n). It is therefore continuous both at these points and between them.

Now, for $r_n \geqslant r \geqslant r_{n+1}$, we may write

$$\frac{x(r)}{r} = \frac{x_n(r-r_{n+1})+x_{n+1}(r_n-r)}{r_n(r-r_{n+1})+r_{n+1}(r_n-r)},$$

which lies between x_n/r_n, x_{n+1}/r_{n+1}, i.e. between $\cos\theta_n$, $\cos\theta_{n+1}$, since $r-r_{n+1}$, r_n-r are both positive. Hence

$$\left|\frac{x(r)}{r}-\cos\alpha_0\right| < \epsilon, \quad \text{if } |\theta_n-\alpha_0| < \text{some } \eta(\epsilon),$$

i.e. if $n >$ some $N(\epsilon)$, by the definition of θ_n,

i.e. if $r_{n+1} <$ some $\rho(\epsilon)$, since $\{r_n\}$ is monotonic,

i.e. if $r < \rho(\epsilon)$, since $r \geqslant r_{n+1}$.

In other words, as $r \to 0$, $x/r \to \cos\alpha_0$, and similarly $y/r \to \sin\alpha_0$. We may complete the definition of the curve by writing

$$x(0) = 0, \qquad y(0) = 0.$$

Then the curve passes through O and Dx, Dy both exist there. Actually at O

$$Dx = \cos\alpha_0, \qquad Dy = \sin\alpha_0,$$

and therefore Dx, Dy do not both vanish at O.

To define a curve in which the umbral derivatives of the *second* order exist at O consider the sequence of points $\{(X_n, Y_n)\}$ where

$$X_n \equiv \frac{x_n}{r_n}-\cos\alpha_0, \qquad Y_n \equiv \frac{y_n}{r_n}-\sin\alpha_0.$$

As we have seen, this sequence converges to the point O and therefore, by the result just established, we can define a continuous curve

$$X = X(r), \qquad Y = Y(r)$$

which passes through a subsequence of these points and O itself, and has derivatives DX, DY at O. Now consider the curve defined by

$$x(r) \equiv r(X+\cos\alpha_0), \qquad y(r) \equiv r(Y+\sin\alpha_0). \tag{17}$$

It passes through a subsequence of the points $\{(x_n, y_n)\}$, i.e. through

a subsequence of the original sequence of points, and is continuous, since $X(r)$, $Y(r)$ are continuous. It evidently passes through O (for $r = 0$) and at O
$$Dx = \cos\alpha_0, \qquad Dy = \sin\alpha_0,$$
so that they cannot both vanish. For the umbral derivatives of the second order we must consider the limits, as $r \to 0$, of
$$\frac{x(r)-x(0)-r\cos\alpha_0}{r^2/2!}, \qquad \frac{y(r)-y(0)-r\sin\alpha_0}{r^2/2!},$$
i.e. of
$$\frac{2X(r)}{r}, \qquad \frac{2Y(r)}{r},$$
which exist, since DX, DY exist at O. Hence the curve $x(r), y(r)$ defined in (17) satisfies all the requirements of the lemma, and the lemma is therefore established. It is evidently not difficult by proceeding inductively in the above way to define a curve in which the umbral derivatives of any specified order exist at O.

6. The function of two variables. Analytical criteria

When we seek to obtain for functions of many variables analytical criteria of a minimum analogous to those obtained in (1), (2) above for functions of one variable, we are confronted with difficulties of a new order. We can isolate one of these difficulties by considering first of all the function of two variables only.

Let the function be $f(x,y)$ supposed differentiable to any required order. Then, by (9), a minimum must also be a stationary point and therefore for a minimum we must have the preliminary conditions
$$f_x = 0 = f_y.$$
Now let us study the function for a minimum at (x,y) along an arbitrary path through the point. The stationary point, as we have seen in (8), is also a stationary point along every differentiable curve through the point. If along the curve the stationary point is also to be a minimum, we have from (2) the conditions
$$[D^2 f] > 0 \quad \text{(sufficient)}, \tag{18}$$
$$[D^2 f] \geq 0 \quad \text{(necessary)}, \tag{19}$$
if $[D^2 f]$ exist at the point.

If Dx, Dy, $[D^2x]$, $[D^2y]$ all exist at the point and if f is twice differentiable, then, by chapter VI (42), $[D^2 f]$ exists at the point, and in fact we have
$$[D^2 f] = f_x[D^2x]+f_y[D^2y]+f_{xx}(Dx)^2+2f_{xy}\,DxDy+f_{yy}(Dy)^2.$$

Thus, at a stationary point of f,

$$[D^2 f] = f_{xx}(Dx)^2 + 2f_{xy}\,DxDy + f_{yy}(Dy)^2. \tag{20}$$

The expression on the right is quadratic in Dx, Dy and therefore, except for Dx, Dy simultaneously zero, has always the sign of f_{xx}, if

$$\begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix} > 0.$$

Hence, if at the point we have simultaneously

$$f_{xx} > 0, \qquad \begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix} > 0, \tag{21}$$

then along every curve through the point at which Dx, Dy exist and are not both zero we have

$$[D^2 f] > 0.$$

Thus, by (18), the conditions (21) are *sufficient* for a minimum.

Conversely, if

$$\begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix} < 0,$$

the quadratic (20) has real factors of the form

$$(\alpha\,Dx + \beta\,Dy)(\alpha' Dx + \beta' Dy),$$

and by a suitable choice of the ratio $Dy : Dx$, i.e. of the tangential direction of the curve, can be given either sign at discretion. Hence, if the point is to be a minimum along *every* differentiable curve through the point and in particular along every straight line through the point, we must have

$$\begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix} \geqslant 0.$$

If this condition is satisfied, the quadratic has the sign of f_{xx} or is zero. Hence for a minimum we have the further necessary condition

$$f_{xx} \geqslant 0.$$

Collecting these results we can state that

(22) *Conditions that a stationary point of $f(x, y)$, supposed differentiable, be also a minimum are*

$$f_{xx} > 0, \quad \begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix} > 0 \qquad (\textit{sufficient});$$

$$f_{xx} \geqslant 0, \quad \begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix} \geqslant 0 \qquad (\textit{necessary}).$$

7. The semi-definite case

The foregoing conditions fail, of course, to provide an exact set of necessary and sufficient conditions for a minimum in view of the gap between '$>$' and '$\geqslant$'. Something of this sort was to be expected, since,

even for a function of one variable only, we were unable to establish an exact condition for a minimum in terms of the second derivative alone. Now

$$f_{xx} = 0, \qquad f_{xx}f_{yy} - f_{xy}^2 > 0$$

are clearly incompatible, and the only divergence between the necessary and the sufficient conditions must come from the possibility

$$\begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix} = 0. \tag{23}$$

In such a case, since also $f_{xx} \geqslant 0$, we may write

$$f_{xx} = \alpha^2, \qquad f_{xy} = \alpha\beta, \qquad f_{yy} = \beta^2,$$

obtaining at the stationary point

$$[D^2f] = (\alpha\, Dx + \beta\, Dy)^2.$$

Thus $[D^2f]$ is still positive along every curve excepting only those which have the unique direction given by

$$\alpha\, Dx + \beta\, Dy = 0.$$

Along every curve in that particular direction $[D^2f]$ vanishes and the decision, by (2), passes to derivatives of higher order. In strictness, it should of course be added that, in the more special case

$$f_{xx} = 0, \qquad f_{xy} = 0, \qquad f_{yy} = 0$$

still covered by (23), $[D^2f]$ vanishes in every direction.

We throw light on these conditions, if we interpret them geometrically. Regard

$$z = f(x,y)$$

as defining a surface† in terms of rectangular cartesians x, y, z in which z measures height and x, y are horizontal coordinates. On a map of this surface the curves of the family

$$f(x,y) = \text{constant}$$

are the contour lines. From the theory of singular points of a plane curve we know that

$$f_x = 0 = f_y$$

are the conditions for a singular point, which will be an isolated (conjugate) point, a real node, a cusp according as

$$\begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix} > 0, \quad < 0, \quad = 0.$$

An isolated point on a contour line, or in the extreme case a contour line which has shrunk to a single point, represents a mountain summit or possibly the pit of some depression: it is thus either a maximum or

† Let us say, some portion of the earth's surface sufficiently small for the level surfaces to be taken as plane.

a minimum. If we think of the surface itself flooded up to the level of the mountain summit, the fact of the isolated point becomes obvious.†

On the other hand, at a real node the contour line crosses itself. The two branches divide the vicinity of the point into four regions, in two of which the surface is above and in the other two below the level of the point. The node on the contour line therefore represents a saddle-back, i.e. the summit of a mountain pass, which is evidently neither a maximum nor a minimum. It is still, however, a stationary point, i.e. the surface is horizontal in its neighbourhood. Geometrically, we may regard it as the two-dimensional analogue of the inflexional case in the function of one variable, where again we have a stationary point which is not a turning-point but has in its immediate vicinity both higher and lower points. I shall extend the term 'inflexional' to all similar points in fields of many variables.

To this order of approximation the cusp lies indecisively between the cases of the isolated point and the real node, and at this stage yields no useful geometrical illustration. We might conveniently apply the term 'cuspidal' to the case in which

$$\begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix} = 0.$$

It is more natural, however, to recall that in this case the quadratic

$$D^2f = f_{xx}(Dx)^2 + 2f_{xy}\,DxDy + f_{yy}(Dy)^2$$

is said, in algebraic language, to be *semi-definite* in the sense that, although it never changes sign for real values of Dx, Dy, yet it does actually attain the value zero. For this reason, then, the case in question will be called the 'semi-definite' case and the term will be similarly applied for functions of more than two variables.

The proper resolution of this semi-definite case is laborious, and it is better to postpone it until we have extended the results of § 6 to functions of many variables.

8. Functions of many variables. The invariably positive quadratic

At a stationary point of $f(x_1,..., x_n)$, supposed twice differentiable, we have

$$[D^2f] = \sum_{r=0}^{n} \frac{\partial^2 f}{\partial x_r^2}(Dx_r)^2 + 2\sum_{r<s}^{n} \frac{\partial^2 f}{\partial x_r\,\partial x_s} Dx_r Dx_s. \tag{24}$$

The expression on the right is a quadratic form in the n arguments $Dx_1,..., Dx_n$. The determination of its sign, or more precisely the

† Compare the point of view of Noah on Ararat.

enumeration of conditions that for all real values of its arguments it be invariably positive or invariably non-negative, belongs properly to the algebra of quadratic forms. In lieu, however, of quoting known results from this algebra, which is by no means elementary, I prefer to obtain the requisite facts as a part of the present theory. It is possible to do this inductively *pari passu* with the development of the conditions for minima of the general function.

The discussion is assisted, if we lay down certain terminology for the general quadratic

$$\sum_{r\ 0}^{n} a_{rr} x_r^2 + 2\sum_{r<s}^{n} a_{rs} x_r x_s.$$

The discriminant of the general quadratic is of course the determinant (here defined by its principal diagonal)

$$|a_{11} \quad a_{22} \quad \dots \quad a_{nn}|.$$

A minor of a term on its principal diagonal such as

$$|a_{22} \quad \dots \quad a_{nn}|$$

will be called a *principal minor* of the determinant. A principal minor of a principal minor will be called a *second principal minor* and so on. By a *sequence of principal minors* I shall mean a set of principal minors such as

$$a_{11},\ |a_{11} \quad a_{22}|,\ \dots,\ |a_{11} \quad a_{22} \quad \dots \quad a_{n-1,n-1}|$$

beginning with a determinant of unit order, in which each determinant is a principal minor of the determinant which follows.

For invariance of sign of the n-ary quadratic we rely on the theorem:

(25) *The necessary and sufficient condition that the n-ary quadratic be invariably positive (except for every argument zero) is that the discriminant and a sequence of principal minors be all positive.*

This theorem applied to the quadratic for $[D^2f]$ in (24) gives a *sufficient* condition that $[D^2f] > 0$ along every curve through the stationary point such that $Dx_1,\dots,Dx_n$ exist and are not all zero at the point. These conditions are sufficient to secure that $f(x_1,\dots,x_n)$ be a minimum at the stationary point. The discriminant of the quadratic for $[D^2f]$ in (24) is the Hessian of f,

$$\left|\frac{\partial^2 f}{\partial x_1^2} \quad \frac{\partial^2 f}{\partial x_2^2} \quad \cdots \quad \frac{\partial^2 f}{\partial x_n^2}\right| \tag{26}$$

Hence, if we have proved (25), we can at once deduce the following conditions sufficient for a stationary point to be a minimum:

(27) *A sufficient condition that $f(x_1,\dots,x_n)$ be minimum at a stationary point is that there the Hessian and a sequence of principal minors be positive.*

9. Proof of the induction

When $n = 2$, the theorem (27) is effectively (22), which we have already proved, while (25) reduces to an elementary property of the binary quadratic which we have actually used in proving (22). The induction therefore opens at $n = 2$ and is sufficiently established when we have shown that (25), (27) hold for n variables, if they hold for $n-1$.

The principles of the proof are sufficiently shown by taking $n = 3$. For the ternary quadratic it is convenient to change the notation and to write
$$u(x,y,z) \equiv (a,b,c,f,g,h)(x,y,z)^2.$$
To prove the conditions of (25) necessary put first of all $z = 0$. The n-ary quadratic reduces to the $(n-1)$-ary quadratic
$$u(x,y,0) = ax^2+2hxy+by^2.$$
Its discriminant $|a \;\; b|$ is a principal minor of $|a \;\; b \;\; c|$, the discriminant of $u(x,y,z)$. Since, by hypothesis, $u(x,y,0)$ is invariably positive, unless $x = 0 = y$, it follows from (25), now supposed to hold for $n-1$ variables, that $|a \;\; b|$ and a principal minor a are positive.

Now give z a fixed value differing from zero and consider variation of x, y only. The set of values $x = 0$, $y = 0$, $z = 0$ is then excluded and therefore, by hypothesis, $u(x,y,z)$ is always positive. It is consequently bounded below and so has a least value u_0 which is positive and is, moreover, a minimum, since x, y are now unrestricted. At this minimum we must have
$$u_x = 0 = u_y,$$
and, since identically
$$u = \tfrac{1}{2}(xu_x+yu_y+zu_z),$$
we have for the minimum value
$$u_0 = \tfrac{1}{2}zu_z.$$
Thus, if (x_0, y_0) be the coordinates of the minimum, we have
$$\left.\begin{aligned} ax_0+hy_0+gz &= 0 \\ hx_0+by_0+fz &= 0 \end{aligned}\right\} \tag{28}$$
and
$$gx_0+fy_0+\left(c-\frac{u_0}{z^2}\right)z = 0.$$
Hence
$$\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c-\dfrac{u_0}{z^2} \end{vmatrix} = 0$$
and so
$$|a \;\; b|u_0 = z^2|a \;\; b \;\; c|. \tag{29}$$

But u_0 is positive and we have already proved that $|a\ \ b| > 0$. Thus also $|a\ \ b\ \ c| > 0$, i.e. the discriminant of $u(x, y, z)$ is positive: and we have shown that a, $|a\ \ b|$ are positive. The conditions of (25) are thus proved *necessary*.

Conversely, to prove the conditions *sufficient*, consider first $z = 0$. Then, as we have seen, (25) reduces to the corresponding theorem in $n-1$ variables, which we have assumed for the purposes of the induction. We may therefore take z not to be zero. Fix it, and consider variation in x, y only. Since $|a\ \ b| \neq 0$ by hypothesis, we can solve equations (29) to give the stationary point (x_0, y_0). The conditions of (27), assumed, in the induction, to hold for $n-1$ variables, are sufficient to ensure that the stationary point be a minimum. But now

$$u_{xx}, u_{xy}, u_{yy} = 2a, 2h, 2b$$

and accordingly the Hessian in x, y of $u(x, y, z)$ differs only by a constant numerical factor from the discriminant of $u(x, y, 0)$.

The Hessian and a sequence of principal minors are thus a sequence of principal minors of $u(x, y, z)$ and so, by hypothesis, are positive. Thus (x_0, y_0) is a minimum of u. Moreover, since $|a\ \ b|$, $|a\ \ b\ \ c|$ are both positive, the minimum value u_0 given by (29) is also positive and hence $u(x, y, z)$ is invariably positive, if $z \neq 0$. This concludes the proof of the sufficiency of (25) and also completes the induction, since, as we have said, the sufficiency of (27) is at once deduced by applying (25) to the expression of $[D^2 f]$ as a quadratic given in (24).

In (25), invariance of sign of the quadratic, a property evidently symmetrical in the arguments, has been made to depend on the signs of a set of expressions which in themselves are not formally symmetrical, namely the discriminant and a set of principal minors. We have actually chosen

$$a > 0, \qquad |a\ \ b| > 0, \qquad |a\ \ b\ \ c| > 0, \tag{30}$$

which display evident alphabetical preferences. Since the conditions of (25) are both necessary and sufficient, it is clear that the three inequalities (30) necessitate the other inequalities

$$b > 0, \qquad c > 0, \qquad |b\ \ c| > 0, \qquad |c\ \ a| > 0$$

which complete the symmetrical set.† Thus it is actually necessary that the discriminant and *all* its principal minors (of every order) be positive. This complete set of conditions is obviously sufficient, since it includes a set of conditions known to be sufficient, but is by that very fact redundant.

† Direct algebraic proof of this is, of course, not difficult.

10. The invariably non-negative quadratic

We can now deduce from (25) the following necessary and sufficient conditions that a quadratic be never negative:

(31) *An n-ary quadratic is invariably positive or zero, if its discriminant and all its principal minors (of every order) are positive or zero.*

As before, I give the proof only for a ternary quadratic $u(x,y,z)$. Define a new ternary quadratic

$$u'(x,y,z) \equiv (a+\epsilon, b+\epsilon, c+\epsilon, f, g, h)(x,y,z)^2 = u+\epsilon(x^2+y^2+z^2),$$

where ϵ is positive but is otherwise at our disposal. Since u is never negative, u' is invariably positive except when x, y, z are all zero. With the usual notation, in which Δ is the discriminant of u and A, B, C its principal minors, the discriminant of u' is

$$\Delta+\epsilon(A+B+C)+\epsilon^2(a+b+c)+\epsilon^3 \tag{32}$$

and a first and a second principal minor are respectively

$$A+\epsilon(b+c)+\epsilon^2, \qquad a+\epsilon. \tag{33}$$

These are all positive and their limits, as $\epsilon \to 0$, are accordingly all positive or zero, i.e.

$$\Delta \geqslant 0, \qquad A \geqslant 0, \qquad a \geqslant 0,$$

and so for the other principal minors. The conditions are therefore necessary.

Conversely, if Δ, A, B, C, a, b, c are all positive or zero, it follows from consideration of (32), (33), since ϵ is positive, that the discriminant of u' and all its principal minors are positive. Hence, by (29), u' is invariably positive. Taking the limit $\epsilon \to 0$, we have that u itself is invariably positive or zero.

Within the terms of (31) it might be supposed that each principal minor and the discriminant itself could, independently of one another, be either positive or zero. Actually not all of these alternatives are algebraically compatible. Suppose, for instance, that $|a\ b\ c|$ be positive. Within the terms of (31) the quadratic, being always positive or zero, is still bounded below and the arguments of § 9 continue to apply. Thus, if u_0 be the minimum, when z is fixed and not zero, (29) shows us that $|a\ b|$ is not zero. Again, as we have already seen, any principal minor (of whatever order) of the discriminant becomes itself the discriminant of the quadratic, when appropriate arguments are equated to zero. It follows that, if the discriminant itself or some principal minor is positive, every principal minor (of whatever order) of the discriminant or of that particular principal minor must also be positive. Thus in any

sequence of principal minors, if any term is positive, every preceding term must also be positive.

In particular, if the discriminant itself be positive, *every* sequence of principal minors is positive and (25) applies. In this case the quadratic is positive and never zero except when every argument is zero. Thus, if under (31) the quadratic actually attains the value zero for other than all zero values of its arguments, the discriminant vanishes. Conversely, if under (31) the discriminant vanishes, the quadratic must actually attain the value zero for other than all zero values of its arguments, since otherwise, by (25), the discriminant is positive.

In such case the zero itself is the minimum and we have for the coordinates of the zero

$$u_x = 0 = u_y, \qquad u_0 = 0,$$

i.e. simultaneously

$$\left.\begin{aligned} ax+hy+gz &= 0 \\ hx+by+fz &= 0 \\ gx+fy+cz &= 0 \end{aligned}\right\}.$$

These are consistent, since the discriminant vanishes. In general they determine the ratios of x, y, z uniquely, so that in homogeneous coordinates (x, y, z) there is a single zero. It may, however, happen that every first minor of the discriminant vanishes and the three equations are equivalent to a single equation. There is then in homogeneous coordinates a line of zeros; in fact, the quadratic is a perfect square $(x\sqrt{a}+y\sqrt{b}+z\sqrt{c})^2$ which vanishes along the line of zeros

$$x\sqrt{a}+y\sqrt{b}+z\sqrt{c} = 0.$$

11. Necessary conditions for a minimum

If we apply (31) with the above glosses to the quadratic for $[D^2f]$, we have the following *necessary* conditions for a minimum:

(34) *If a stationary point of $f(x_1,..., x_n)$ is also a minimum, it is necessary that the Hessian and every principal minor (of whatever order) be either positive or zero. Moreover, any of these which is positive must have every principal minor also positive.*

As we have seen, if, within the terms of (31), the Hessian is positive, necessarily every principal minor is also positive, the conditions of (25) are satisfied and the point is actually a minimum. Cases of failure of the tests of (25) therefore arise only if the Hessian vanishes. For two variables this is the case $f_{xx}f_{yy}-f_{xy}^2 = 0$, which is exactly the semi-definite or 'cuspidal' case. We may accordingly in any number

of variables describe the case in which the Hessian vanishes as the 'semi-definite' case. In this semi-definite case $[D^2f]$ vanishes along curves in certain directions through the stationary point, and along such curves we have to consider derivatives of higher order.

Actually these exceptional directions are given by the n equations

$$f_{r1}\,Dx_1+\ldots+f_{rn}\,Dx_n = 0 \qquad (r = 1,\ldots,n),$$

which are homogeneous in Dx_r and are consistent, since the Hessian vanishes. They may define a unique direction through the stationary point, or they may be equivalent to not more than $n-s$ independent equations and define an $(s-1)$-fold of directions. The discussion of this semi-definite case we postpone to the next chapter.

12. Minimum distance from a quadric

To show the application of these tests in practice we may inquire in what circumstances the normal distance of a point from a quadric surface is actually a maximum or minimum distance of the point from the surface. We take as coordinate-axes the normal and the directions of principal curvature at a point O. The quadric then takes the form

$$2z = x^2/\rho_1+y^2/\rho_2 \qquad (\rho_1, \rho_2 \neq 0), \tag{35}$$

where ρ_1, ρ_2 are the principal radii of curvature at O. Take the fixed point N on the normal at O as $(0, 0, \zeta)$. Then, if P be the point (x, y, z) of the quadric, we have

$$\begin{aligned} f(x,y,z) \equiv PN^2 &= x^2+y^2+(\zeta-z)^2 \\ &= \zeta^2+x^2(1-\zeta/\rho_1)+y^2(1-\zeta/\rho_2)+\tfrac{1}{4}(x^2/\rho_1+y^2/\rho_2)^2, \end{aligned}$$

on substitution for z from (35). Clearly O itself gives a stationary point, and at O

$$D^2f = (1-\zeta/\rho_1)(Dx)^2+(1-\zeta/\rho_2)(Dy)^2. \tag{36}$$

By (27), conditions sufficient for a minimum are

$$1-\zeta/\rho_1 > 0, \qquad (1-\zeta/\rho_1)(1-\zeta/\rho_2) > 0,$$

i.e.
$$\rho_1(\rho_1-\zeta) > 0, \qquad \rho_2(\rho_2-\zeta) > 0.$$

These are the conditions that C_1, C_2, the principal centres of curvature, do not lie between O, N. Similarly, it is sufficient for a maximum that

$$\rho_1(\rho_1-\zeta) < 0, \qquad \rho_2(\rho_2-\zeta) < 0,$$

i.e. that C_1, C_2 both lie between O, N. If of C_1, C_2 one lies between O, N and the other outside them, that is to say, if the segments C_1C_2, ON overlap, then

$$(1-\zeta/\rho_1)(1-\zeta/\rho_2) < 0$$

and the stationary position is inflexional.

There remains to be considered the semi-definite case in which

$$(1-\zeta/\rho_1)(1-\zeta/\rho_2) = 0,$$

i.e. N coincides with C_1 or C_2, say with C_1. For a minimum we still need

$$1-\zeta/\rho_2 \geqslant 0, \tag{37}$$

i.e. C_2 does not lie between O, N but may coincide with N, i.e. C_1, C_2 may coincide. With this condition satisfied we have in the semi-definite case

$$D^2 f \geqslant 0$$

and similarly from (37)

$$x^2(1-\zeta/\rho_1)+y^2(1-\zeta/\rho_2) \geqslant 0.$$

Hence

$$\delta f \geqslant \tfrac{1}{4}(x^2/\rho_1+y^2/\rho_2)^2 \geqslant 0.$$

These conditions therefore suffice to secure a minimum.

Again, when $\zeta = \rho_1$, if we put $x \neq 0$, $y = 0$, we have

$$\delta f = \tfrac{1}{4}(x^2/\rho_1)^2 > 0.$$

The semi-definite case is thus incompatible with a maximum, which requires $\delta f \leqslant 0$. Hence, if C_2 lies between O, N, the position is inflexional.

We may sum up these conclusions as follows:

(38) *If N is a fixed point on the normal at O to a quadric, the normal distance NO is a minimum distance from N to the quadric, if neither centre C_1, C_2 of principal curvature at O lies between N, O. The normal distance is a maximum, if C_1, C_2 both lie between N, O. If the segments C_1C_2, NO overlap, the normal distance is inflexional.*

When N is at C_1, say, the normal distance is a minimum, if C_2 does not lie between O, C_1 or is at C_1 itself. It is inflexional, if C_2 lies between O, C_1.

13. Minimum distance between two quadrics

As an example of the determination of maxima and minima of a function of four arguments we may treat the analogous problem of the distance between points of the two quadrics

$$2z = x^2/\rho_1+y^2/\rho_2 \qquad (\rho_1, \rho_2 \neq 0),$$
$$2z = 2\zeta+x^2/\rho_1'+y^2/\rho_2' \qquad (\rho_1', \rho_2' \neq 0)$$

in the neighbourhood of their common normal, the z-axis. If P, P' be the points (x, y, z), (x', y', z') of the two quadrics, we have

$$\begin{aligned} f(x,x',y,y',z,z') = PP'^2 &= (x-x')^2+(y-y')^2+(z-z')^2 \\ &= \zeta^2+x^2(1-\zeta/\rho_1)+x'^2(1+\zeta/\rho_1')-2xx'+ \\ &\quad +y^2(1-\zeta/\rho_2)+y'^2(1+\zeta/\rho_2')-2yy'+ \\ &\quad +\tfrac{1}{4}(x^2/\rho_1+y^2/\rho_2-x'^2/\rho_1'-y'^2/\rho_2')^2. \end{aligned}$$

The common normal gives, of course, a stationary value of PP', and in its vicinity we have

$$D^2f = (1-\zeta/\rho_1)(Dx)^2+(1+\zeta/\rho_1')(Dx')^2-2DxDx'+ \\ +(1-\zeta/\rho_2)(Dy)^2+(1+\zeta/\rho_2')(Dy')^2-2DyDy'.$$

The discriminant of D^2f reduces to the product of the determinants

$$\begin{vmatrix} 1-\zeta/\rho_1 & -1 \\ -1 & 1+\zeta/\rho_1' \end{vmatrix}, \qquad \begin{vmatrix} 1-\zeta/\rho_2 & -1 \\ -1 & 1+\zeta/\rho_2' \end{vmatrix}, \tag{39}$$

which are themselves principal minors of the discriminant. The other principal minors are

$$1-\zeta/\rho_1, \qquad 1+\zeta/\rho_1', \qquad 1-\zeta/\rho_2, \qquad 1+\zeta/\rho_2' \tag{40}$$

together with their products by that determinant of (39) which has the opposite suffix.

For a minimum, it is consequently sufficient that the six expressions (39), (40) be all positive, while it is necessary that none be negative. As in the previous example, the necessary conditions secure $D^2f \geqslant 0$, and at the same time

$$x^2(1-\zeta/\rho_1)+x'^2(1+\zeta/\rho_1')-2xx'+y^2(1-\zeta/\rho_2)+y'^2(1+\zeta/\rho_2')-2yy' \geqslant 0,$$

and hence they secure

$$\delta f \geqslant \tfrac{1}{4}(x^2/\rho_1+y^2/\rho_2-x'^2/\rho_1'-y'^2/\rho_2')^2 \geqslant 0.$$

The necessary conditions are thus also sufficient conditions. Consideration of (39), (40) shows that these conditions are symmetrical in the suffixes 1, 2 and that therefore we need consider in detail only those conditions which involve, say, the suffix 1. We have then from (39)

$$\rho_1\rho_1'\,\zeta(\rho_1-\rho_1'-\zeta) \geqslant 0,$$

and from (40) $\qquad \rho_1(\rho_1-\zeta) \geqslant 0, \qquad \rho_1'(\rho_1'+\zeta) \geqslant 0.$

If O, O' are $(0,0,0)$, $(0,0,\zeta)$ the extremities of the common normal, and if C_1, C_1' are $(0,0,\rho_1)$, $(0,0,\zeta+\rho_1')$ principal centres of curvature, we can state these conditions as

$$C_1O\,.\,C_1'O'\,.\,O'O\,.\,C_1C_1' \geqslant 0, \tag{41}$$

$$C_1O'\,.\,C_1O \geqslant 0, \qquad C_1'O\,.\,C_1'O' \geqslant 0. \tag{42}$$

Of these, (41) is the condition that the segments $O'C_1$, OC_1' overlap; (42) is the condition that C_1, C_1' both lie outside OO'. The only possible orders of the four points are thus

$$OO'C_1'C_1 \quad \text{or} \quad OO'C_1C_1'.$$

There also fall to be considered possibilities of coincidence of these

points, corresponding to the semi-definite case. In (41) we then have $O \equiv O'$ or $C_1 \equiv C_1'$.† If O, O' coincide, (42) is evidently satisfied and so from symmetry are the conditions in suffix 2. But the normal distance is now zero and is therefore obviously a minimum. If C_1, C_1' coincide, the two conditions of (42) become identical. They require that the coincident points C_1, C_1' lie outside OO'. Coincidence with either O or O' is, of course, excluded for the reasons of the footnote. Moreover, we cannot have $C_1 O'$, $C_1' O$ overlapping and C_1, O' coincident, for this again requires either C_1' or O to be also coincident with C_1, O'. This confirms what we have already remarked, namely that we cannot have the second order discriminant $|a\ b|$ positive when a principal term a or b is zero.

For a maximum, the minors (39) are still positive, but the minors (40) are all negative, i.e. the inequality (41) is unchanged but the inequalities (42) are reversed. Thus, for a maximum, $C_1 O'$, $C_1' O$ overlap and C_1, C_1' lie both inside OO'. The only possible order of the four points is

$$OC_1 C_1' O'$$

As in the previous example it can be seen that the semi-definite case is incompatible with a maximum and therefore, for a maximum, coincidence of C_1, C_1' etc. must be excluded.

It will be noticed that in the case of either a maximum or a minimum the segments OC_1, $O'C_1'$ do not overlap and we can state a *sufficient* condition for the stationary position to be a turning position, namely that $O'C_1$, OC_1' overlap and OC_1, $O'C_1'$ do not overlap.

We may sum up these conclusions as follows:

(43) *If two quadrics have a common normal OO' and parallel directions of curvature at O, O' and if C_1, C_1', C_2, C_2' be their corresponding centres of principal curvature, then the normal distance OO' is a minimum distance between the two quadrics if the above points lie in the orders*

$$OO'C_1' C_1 \quad or \quad OO'C_1 C_1',$$

and

$$OO'C_2' C_2 \quad or \quad OO'C_2 C_2',$$

where coincidence of any of the pairs O, O', C_1, C_1', C_2, C_2' is permissible.

The distance OO' is a maximum, if the points lie in the orders

$$OC_1 C_1' O', \qquad OC_2 C_2' O'$$

and coincidences are not permitted.

† We do not consider $C_1 = O$ or $C_1 = O'$, since these are expressly excluded in defining the quadrics and would invalidate our analysis. Moreover, these conditions do not give a zero discriminant.

WORKED EXAMPLE

Show that the sum of the products, taken p together, of the squared distances of a point from n fixed points is a minimum at every stationary point, if $n > 3(p-1)$.

For the sum of the products, taken r together, of the squared distances $(x-x_s)^2+(y-y_s)^2+(z-z_s)^2$, when the suffixes $a, b, \ldots$ are omitted from $1, 2, \ldots, n$, let us write

$$\overset{r}{\underset{(a,b,\ldots)}{S}}.$$

Then the function to be tested is

$$u(x,y,z) = \overset{p}{S}$$

and we have

$$u_x = 2\sum_s (x-x_s)\overset{p-1}{\underset{(s)}{S}},$$

$$u_{xx} = 2(n-p+1)\overset{p-1}{S} + 8\sum_{s,t}(x-x_s)(x-x_t)\overset{p-2}{\underset{(s,t)}{S}},$$

$$u_{xy} = 4\sum_{s,t}(x-x_s)(y-y_t)\overset{p-2}{\underset{(s,t)}{S}}.$$

Choose any real numbers ξ, η, ζ and write

$$\theta_s = \xi(x-x_s)+\eta(y-y_s)+\zeta(z-z_s).$$

Then $$(\xi\partial_x+\eta\partial_y+\zeta\partial_z)^2 u = 2(n-p+1)(\xi^2+\eta^2+\zeta^2)\overset{p-1}{S} + 8\sum_{s,t}\theta_s\theta_t\overset{p-2}{\underset{(s,t)}{S}}.$$

Now in the last term on the right pick out from S the particular product

$$\prod_{s=1}^{p-2}\{(x-x_s)^2+(y-y_s)^2+(z-z_s)^2\}.$$

In $\sum_{s,t}\theta_s\theta_t\overset{p-2}{\underset{(s,t)}{S}}$ this particular product is multiplied by

$$\sum_{s,t=p-1}^{n} 8\theta_s\theta_t = 4(\theta_{p-1}+\ldots+\theta_n)^2-4(\theta_{p-1}^2+\ldots+\theta_n^2)$$

$$\geq -4(\xi^2+\eta^2+\zeta^2)\sum_{s=p-1}^{n}\{(x-x_s)^2+(y-y_s)^2+(z-z_s)^2\},$$

since $$\theta_s^2 < (\xi^2+\eta^2+\zeta^2)\{(x-x_s)^2+(y-y_s)^2+(z-z_s)^2\}.$$

Thus $$(\xi\partial_x+\eta\partial_y+\zeta\partial_z)^2 u > 2(\xi^2+\eta^2+\zeta^2)\left\{(n-p+1)\overset{p-1}{S}-2(p-1)\overset{p-1}{S}\right\}$$

$$> 0, \quad \text{if} \quad n > 3(p-1).$$

Since ξ, η, ζ are any real numbers, this condition secures that $D^2u > 0$ at every stationary point, which is therefore a minimum.

EXAMPLES IX

1. Prove that a point at which $D^r f(x) = 0$ $(r = 1, 2, \ldots, 2n)$ is a minimum of $f(x)$, if it is a minimum of any $D^{2s}f(x)$ $(s = 1, 2, \ldots, n)$. Prove also the converse, if $f(x)$ is sufficiently differentiable; but by considering the functions

$$x^4\sin^2(x^{-1}), \qquad x^4\sin^2\log|x|,$$

or otherwise, show that the converse is not true in general.

2. If a, b, p, q are unequal, show that

$$f(x) = \frac{p^2}{x-a}-\frac{q^2}{x-b}$$

has a single maximum and a single minimum, and that the function does not lie between its maximum and its minimum. Confirm this algebraically.

3. Determine the turning values (if any) of

$$\frac{a+b\cos x}{a'+b'\cos x},\qquad \frac{a\cos x+b\sin x}{a'\cos x+b'\sin x},\qquad \frac{a^2\cos^2 x+b^2\sin^2 x}{a'^2\cos^2 x+b'^2\sin^2 x},$$

$$(ax^2+x+1)e^x,\qquad (ax^2+x+a)e^x,$$

$$\frac{\sin^2 x}{x^2},\qquad \frac{\sin^2 kx}{\sin^2 x}\quad (k \text{ not an integer}).$$

4. Show that $$\left(1+\frac{a}{x}\right)^x,\qquad \frac{a^x-1}{x}$$

have neither maxima nor minima.

5. Determine the maxima and minima in the interval $(0, 2\pi)$ of the finite sum

$$\sum_{r=k+1}^{k+n} \frac{1}{r}\cos(rx+\alpha),$$

where k need not be an integer.

6. Determine the maxima and minima in the interval $(0,\pi)$ of the finite sum

$$\sin x+\frac{\sin 2x}{2}+\dots+\frac{\sin nx}{n},$$

and show that, as we pass from 0 to π, the values at the successive maxima steadily decrease.

Show that the greatest value of the sum in $(0,\pi)$, regarded as a function of n, increases steadily with n.

7. Determine the maxima and minima in the interval $(0,\pi)$ of the finite sum

$$\cos x+\frac{\cos 2x}{2}+\dots+\frac{\cos nx}{n},$$

and show that, as we pass from 0 to π, the values at successive maxima and at successive minima both steadily increase.

Obtain the greatest and the least value of the given sum in the given interval.

8. Determine the maxima and minima in the interval $(0,\pi)$ of the finite sum

$$\tfrac{1}{2}x+\sin x+\frac{\sin 2x}{2}+\dots+\frac{\sin nx}{n},$$

and show that, as we pass from 0 to π, the values at the maxima decrease and those at the minima increase.

Obtain the greatest and least values of the sum in the interval $(0, 2\pi)$.

9. Show that the stationary points of

$$a_1\sin\theta\sin\phi+a_2\sin\theta\cos\phi+a_3\cos\theta\sin\phi+a_4\cos\theta\cos\phi$$

are given by

$$\tan 2\theta=-\frac{2(a_1a_3+a_2a_4)}{a_1^2+a_2^2-a_3^2-a_4^2},\qquad \tan 2\phi=-\frac{2(a_1a_2+a_3a_4)}{a_1^2-a_2^2+a_3^2-a_4^2},$$

and determine which (if any) give maxima and minima.

10. (i) Show that $$(ax^2+2bxy+cy^2)e^x$$

has two stationary points, one or other of which is a turning point according as $b^2 \gtrless ac$.

(ii) Show that $$x\exp\{-(ax^2+2bxy+cy^2)\}$$

has neither maximum nor minimum, if either of c, $ac-b^2$ is negative, and determine its maxima and minima, if these expressions are both positive.

Discuss the maxima and minima of the functions (i), (ii), when $ac = b^2$.

11. If u, U denote the general linear and general quadratic form in n variables $(x_1, .., x_n)$, show that after a suitable homographic transformation we can write

$$U = a_1 x_1^2 + a_2 x_2^2 + \ldots + a_n x_n^2, \qquad u = x_1.$$

Investigate the maxima and minima of the functions

$$(a_1 x_1^2 + \ldots + a_n x_n^2)\exp(x_1), \qquad x_1 \exp\{-(a_1 x_1^2 + \ldots + a_n x_n^2)\},$$
$$(a_1 x_1^2 + \ldots + a_n x_n^2)\exp\{-(b_1 x_1^2 + \ldots + b_n x_n^2)\}.$$

12 Investigate the stationary and the turning values of

$$\sum_{r=1}^{n} x_r^3 - 3\sum_{r>s}^{n} x_r x_s.$$

13. If $p_1, .., p_n$ denote the sum of the products $1, \ldots, n$ at a time of the n arguments $x_1, .., x_n$ and if $A_1, \ldots, A_n$ are any constants, show that the expression

$$A_1 p_1 + A_2 p_2 + \ldots + A_n p_n,$$

has no turning values.

14. Show that an implicit function $z(x, y)$ defined by an equation of the form

$$xy + x f(z) + y g(z) + h(z) = 0$$

can have neither maxima nor minima.

15. If u, v are homogeneous quadratic functions of n variables $x_1, \ldots, x_n$, show that the stationary values of u/v are the zeros of the discriminant of $u - \lambda v$, and that the maximum and the minimum of u/v are respectively the greatest and the least zeros of this discriminant.

16. Show that the stationary values of

$$u = \frac{ax^3 + by^3 + cz^3}{3(y+z)(z+x)(x+y)}$$

are the roots of $$\pm\sqrt{\left(\frac{u}{u+a}\right)} \pm \sqrt{\left(\frac{u}{u+b}\right)} \pm \sqrt{\left(\frac{u}{u+c}\right)} = 1,$$

the corresponding values of x, y, z being proportional to

$$\pm\sqrt{\left(\frac{u}{u+a}\right)}, \quad \pm\sqrt{\left(\frac{u}{u+b}\right)}, \quad \pm\sqrt{\left(\frac{u}{u+c}\right)}$$

Investigate the maxima and minima of the function.

17. A function $z(x, y)$ satisfies a partial differential equation of the form

$$\alpha z_{xx} + 2\beta z_{xy} + \gamma z_{yy} = 0,$$

where α, β, γ are constants (or functions of x, y) such that $\alpha\gamma > \beta^2$. Show that in any region the function $z(x, y)$ assumes its greatest values only *on the boundary* of the region.

Extend this result to a function $z(x_1, \ldots, x_n)$ which satisfies a partial differential equation of the second order

$$\sum \alpha_{rr} \frac{\partial^2 z}{\partial x_r^2} + \sum \alpha_{rs} \frac{\partial^2 z}{\partial x_r \partial x_s},$$

where α_{rr}, α_{rs} are the coefficients of an invariably positive quadratic.

18. Given n points in space $P_1, \ldots, P_n$, show that $PP_1^2 + \ldots + PP_n^2$ has a single stationary value, namely when P is at the mean centre of the n points. Show also that this stationary value is the least value of the sum.

19. Given n non-parallel lines in a plane, show that there is a single point of the plane such that the sum of the squares of its distances from the given lines is stationary, and show that the sum is actually a minimum at this point.

Prove analogous results for (i) n given planes, (ii) n given lines in space.

20. Given n points in a plane, determine the positions of a line such that the sum of the squares of its distances from the n points is stationary. Show that there are two such lines which meet at right angles in the mean centre of the n points.

Which (if either) of these two lines gives a maximum or minimum?

21. Show that the product of the distances (taken positively) of a point from n given lines coplanar with it is stationary at the $\frac{1}{2}n(n-1)$ intersections of the n lines, and also at $\frac{1}{2}(n-1)(n-2)$ other points which are the residual intersections of two $(n-1)$-ics passing through all the intersections of the n lines. Show also that the product is a minimum at points of the first set and a maximum at points of the second set, if real.

Obtain an analogous result for the product of the distances of a point from n given planes.

22. Show that the product of the distances (taken positively) of a point from n given points coplanar with it is nowhere a maximum and is a minimum only at the points themselves.

Extend this result (i) to the ratio of two such products, (ii) to the case in which the moving point is no longer in the plane of the given points.

23. *A is one corner of a cube and B, C, D are the corners of the cube nearest to A.* Show that the product

$$PA^2 . PB^2 . PC^2 . PD^2$$

is stationary when P is at O, the centre of the cube; when P is at A, B, C, D; and when P divides one of OA, OB, OC, OD in the ratio of one to three. Show that there are no other stationary points; that O, A, B, C, D alone give minima; and that there are no maxima.

24. With the data of example 23 show that

$$PB^2 . PC^2 . PD^2 + PC^2 . PD^2 . PA^2 + PD^2 . PA^2 . PB^2 + PA^2 . PB^2 . PC^2$$

has no stationary point other than O, and that this point is a minimum.

25. Show that the product of the squared distances of a point from the eight corners of a cube is a minimum at the centre of the cube and at the eight corners themselves.

Show that there are just eight other stationary points (which are the corners of a cube lying within the given cube), but that these are not turning-points.

X

MAXIMA AND MINIMA (*continued*)

1. Restricted minima

It is often necessary to determine the maxima and minima of a function $f(x_1,..., x_n)$ whose arguments $x_1,..., x_n$ are not all independent but are restricted by m relations

$$\phi_1(x_1,..., x_n) = 0, \quad ..., \quad \phi_m(x_1,..., x_n) = 0. \tag{1}$$

Evidently we must have $m < n$, and the function f has in effect just $n-m$ degrees of freedom. Maxima and minima determined under such restrictions may be called 'restricted' maxima and minima.

It is, of course, possible in theory to use the m relations (1) to eliminate m of the arguments from f and so to exhibit f as a function of $n-m$ really independent arguments. But this elimination is frequently impracticable and more often undesirable on grounds of symmetry and convenience.

For simplicity of statement it will be enough to suppose $m = 1$, taking the restrictive condition as

$$\phi(x_1,..., x_n) = 0. \tag{2}$$

It will be found that there is no consequent loss of generality in argument or procedure. I shall, moreover, as before, think only of *minima* of the function.

We have then in the first place to determine the stationary points of the function $f(x_1,..., x_n)$ subject to the restriction (2). At a stationary point we have along an arbitrary differentiable curve

$$f_1 Dx_1 + ... + f_n Dx_n = 0. \tag{3}$$

But this 'curve' must (in geometrical language) 'lie on the surface' (2), i.e. $Dx_1,..., Dx_n$ must satisfy the equation

$$\phi_1 Dx_1 + ... + \phi_n Dx_n = 0. \tag{4}$$

Subject to this restriction, the differentials $Dx_1,..., Dx_n$ are arbitrary and independent. Hence, at a stationary point, (3) must be satisfied for all sets of values of $Dx_1,..., Dx_n$ that satisfy (4).

Multiplying equation (2) by an arbitrary λ and adding to (3), we get

$$(f_1+\lambda\phi_1) Dx_1 + ... + (f_n+\lambda\phi_n) Dx_n = 0. \tag{5}$$

Let us imagine that equation (4) has been used to express one of the arguments, say x_1, in terms of the remaining arguments $x_2, x_3,..., x_n$, so

that $Dx_2, Dx_3, \ldots, Dx_n$ are to be regarded as all independent but Dx_1 as a function of them. Choose the arbitrary λ to satisfy the equation

$$f_1+\lambda\phi_1 = 0.$$

Then (5) becomes

$$(f_2+\lambda\phi_2)\,Dx_2+(f_3+\lambda\phi_3)\,Dx_3+\ldots+(f_n+\lambda\phi_n)\,Dx_n = 0.$$

Since $Dx_2, Dx_3, \ldots, Dx_n$ are now all independent, their coefficients must severally vanish, i.e. we must have the conditions

$$f_2+\lambda\phi_2 = 0, \qquad f_3+\lambda\phi_3 = 0, \quad \ldots, \quad f_n+\lambda\phi_n = 0.$$

In other words, the condition for a stationary point is that at the point it be possible to choose λ to satisfy simultaneously the n equations

$$f_1+\lambda\phi_1 = 0, \quad \ldots, \quad f_n+\lambda\phi_n = 0. \tag{6}$$

The equations determining the stationary point or points may therefore be written in determinant form as

$$\left\| \begin{matrix} f_1 & . & . & . & f_n \\ \phi_1 & . & . & . & \phi_n \end{matrix} \right\| = 0. \tag{7}$$

This condition has been shown to be necessary and it is certainly sufficient, for, if equations (6) be satisfied, evidently (3) is a consequence of (4). It should be remarked that the equations (6) or (7) represent just $n-1$ independent conditions, which, with the condition (2), give the n conditions necessary to determine a definite point or points in the field of the n arguments $x_1,\ldots,x_n$.

We can enunciate the corresponding conditions in the general case when there are m equations of condition ($m < n$):

(8) *If the n arguments $x_1,\ldots,x_n$ are connected by the m relations* ($m < n$)

$$\phi(x_1,\ldots,x_n) = 0, \quad \ldots, \quad \omega(x_1,\ldots,x_n) = 0,$$

then the stationary points of the function $f(x_1,\ldots,x_n)$ are given by the elimination of the m unknowns $\lambda_1,\ldots,\lambda_m$ between the n equations

$$\frac{\partial f}{\partial x_i} = \lambda_1\frac{\partial \phi}{\partial x_i}+\ldots+\lambda_m\frac{\partial \omega}{\partial x_i} = 0, \qquad (i = 1,\ldots,n),$$

i.e. are determined by the set of $n-m$ equations

$$\left\| \begin{matrix} \dfrac{\partial f}{\partial x_1} & . & . & . & \dfrac{\partial f}{\partial x_n} \\ \dfrac{\partial \phi}{\partial x_1} & . & . & . & \dfrac{\partial \phi}{\partial x_n} \\ . & . & . & . & . \\ \dfrac{\partial \omega}{\partial x_1} & . & . & . & \dfrac{\partial \omega}{\partial x_n} \end{matrix} \right\| = 0.$$

2. Discrimination of minima. The restricted quadratic

For the second derivative we have (along any curve for which the derivatives D^2x_s exist)

$$D^2f = \sum f_s\, D^2x_s + \sum f_{ss}\,(Dx_s)^2 + 2\sum f_{st}\, Dx_s\, Dx_t, \tag{9}$$

where in the first two summations s takes all integer values in $(1, n)$ and in the third summation s, t take all different pairs of unequal integer values in $(1, n)$.

Along such a curve we similarly have from the equation of condition

$$0 = D^2\phi = \sum \phi_s\, D^2x_s + \sum \phi_{ss}\,(Dx_s)^2 + 2\sum \phi_{st}\, Dx_s\, Dx_t. \tag{10}$$

At a stationary point, if λ has the value determined by (6), we have, by considering the expression $D^2f+\lambda\, D^2\phi$,

$$D^2f = \sum\,(f_{ss}+\lambda\phi_{ss})(Dx_s)^2 + 2\sum\,(f_{st}+\lambda\phi_{st})Dx_s\, Dx_t. \tag{11}$$

For a minimum, a sufficient condition and a necessary condition are respectively that $D^2f > 0$, and that $D^2f > 0$, for all values of $Dx_1,\ldots, Dx_n$ that satisfy the equations of condition (4).

The problem of determining restricted minima has thus been reduced to a problem in the theory of Quadratic Forms, namely that of obtaining the conditions under which a quadratic

$$u \equiv \sum a_{ss}\, x_s^2 + 2\sum a_{st}\, x_s\, x_t \tag{12}$$

is invariably positive (or is never negative) for all real values of arguments x_s that satisfy the linear relation

$$\sum p_s\, x_s = 0. \tag{13}$$

Let us for simplicity of discussion take $n = 3$. Then in the language of Analytical Geometry we require the conditions that along the line

$$p_1 x + p_2 y + p_3 z = 0$$

the quadratic

$$u \equiv a_{11} x^2 + a_{22} y^2 + a_{33} z^2 + 2a_{23}\, yz + 2a_{13}\, zx + 2a_{12}\, xy$$

be invariably positive (or be never negative). Under these conditions the line cuts the conic $u = 0$ in imaginary points (or else in imaginary or coincident points). By elementary algebra the condition for this is found to be that

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & p_1 \\ a_{12} & a_{22} & a_{23} & p_2 \\ a_{13} & a_{23} & a_{33} & p_3 \\ p_1 & p_2 & p_3 & 0 \end{vmatrix} < 0 \quad \text{or} \quad \leqslant 0. \tag{14}$$

If (14) is satisfied, then u maintains an invariable sign (or never changes sign) along $p_1 x + p_2 y + p_3 z = 0$. To determine this sign we

put $z = 0$; then the sign is positive, if u is positive at the point

$$p_1x+p_2y = 0 = z.$$

The condition for this is that

$$\begin{vmatrix} a_{11} & a_{12} & p_1 \\ a_{12} & a_{22} & p_2 \\ p_1 & p_2 & 0 \end{vmatrix} < 0.$$

3. The restricted Hessian

If we apply this algebraic theory to the quadratic (11) with the restrictive condition (4), still keeping $n = 3$, we obtain the corresponding conditions

$$\begin{vmatrix} f_{11}+\lambda\phi_{11} & f_{12}+\lambda\phi_{12} & f_{13}+\lambda\phi_{13} & \phi_1 \\ f_{12}+\lambda\phi_{12} & f_{22}+\lambda\phi_{22} & f_{23}+\lambda\phi_{23} & \phi_2 \\ f_{13}+\lambda\phi_{13} & f_{23}+\lambda\phi_{23} & f_{33}+\lambda\phi_{33} & \phi_3 \\ \phi_1 & \phi_2 & \phi_3 & 0 \end{vmatrix} < 0 \quad \text{or} \quad \lessgtr 0, \tag{15}$$

$$\begin{vmatrix} f_{11}+\lambda\phi_{11} & f_{12}+\lambda\phi_{12} & \phi_1 \\ f_{12}+\lambda\phi_{12} & f_{22}+\lambda\phi_{22} & \phi_2 \\ \phi_1 & \phi_2 & 0 \end{vmatrix} < 0, \tag{16}$$

where, at the stationary point,

$$f_1+\lambda\phi_1 = f_2+\lambda\phi_2 = f_3+\lambda\phi_3.$$

I shall call (15) the 'restricted' Hessian of f. It is evidently the Hessian of $f+\lambda\phi$, where λ is a constant, bordered by the first derivatives of the 'restrictive' function ϕ. The determinant (16) is a principal first minor of the restricted Hessian, in which the 'bordering' rows and columns have been left intact in forming the minor. By a principal minor of a *restricted* Hessian I shall always understand such a principal minor in which bordering rows and columns are undisturbed. Finally, we observe that the inequalities in (15), (16) have both been reversed. We can check the sign of the inequalities by taking a principal minor of (16), namely

$$\begin{vmatrix} f_{11}+\lambda\phi_{11} & \phi_1 \\ \phi_1 & 0 \end{vmatrix} = -\phi_1^2.$$

If $m = 2$, analogy suggests for the appropriate restricted Hessian

$$\begin{vmatrix} f_{11}+\lambda\phi_{11}+\mu\psi_{11} & f_{12}+\lambda\phi_{12}+\mu\psi_{12} & f_{13}+\lambda\phi_{13}+\mu\psi_{13} & \phi_1 & \psi_1 \\ f_{12}+\lambda\phi_{12}+\mu\psi_{12} & f_{22}+\lambda\phi_{22}+\mu\psi_{22} & f_{23}+\lambda\phi_{23}+\mu\psi_{23} & \phi_2 & \psi_2 \\ f_{13}+\lambda\phi_{13}+\mu\psi_{13} & f_{23}+\lambda\phi_{23}+\mu\psi_{23} & f_{33}+\lambda\phi_{33}+\mu\psi_{33} & \phi_3 & \psi_3 \\ \phi_1 & \phi_2 & \phi_3 & 0 & 0 \\ \psi_1 & \psi_2 & \psi_3 & 0 & 0 \end{vmatrix}.$$

The signs of the inequalities are now given by that of

$$\begin{vmatrix} \times & \times & \phi_2 & \psi_2 \\ \times & \times & \phi_3 & \psi_3 \\ \phi_2 & \phi_3 & 0 & 0 \\ \psi_2 & \psi_3 & 0 & 0 \end{vmatrix},$$

where $\times$ indicates an irrelevant constituent. The determinant reduces to $+(\phi_2\psi_3-\phi_3\psi_2)^2$, and we infer that the restricted Hessian and its principal minors are now positive. More generally, we infer that the restricted Hessian and its principal minors have the sign $(-)^m$, where m is the number of restrictive conditions.

So far this general theory rests only on analogy and surmise. We establish it on a sounder basis by connecting the restricted Hessian with the Hessian of an unrestricted function. For simplicity, again take $m = 1$, $n = 3$, though the arguments will be seen to be perfectly general. Suppose, then, that we have to discuss the function $f(x,y,z)$ subject to the condition $\phi(x,y,z) = 0$. Let us assume that we can solve this equation for z, and that we substitute for z in f and so get

$$F(x,y) = f(x,y,z), \tag{17}$$

where z is regarded as a function of x, y defined by the equation

$$0 = \phi(x,y,z). \tag{18}$$

We are now in a field of two independent variables x, y and shall indicate partial differentiation in these two variables by literal suffixes x, y; we shall retain numerical suffixes 1, 2, 3 to indicate partial differentiation of f, ϕ in their three arguments x, y, z regarded for that purpose as all independent.

Define the function $\lambda(x,y,z)$ by the equation

$$f_3+\lambda\phi_3 = 0. \tag{19}$$

Evidently at the stationary point this *function* λ assumes the value already assigned to the *constant* λ, and so no confusion arises. Differentiating (17), (18) partially in x we get

$$F_x = f_1+f_3 z_x, \qquad 0 = \phi_1+\phi_3 z_x,$$

whence

$$F_x = f_1+\lambda\phi_1,$$

and similarly, by differentiation in y,

$$F_y = f_2+\lambda\phi_2.$$

Differentiating these partially in x and in y we get

$$F_{xx} = f_{11}+\lambda\phi_{11}+(f_{13}+\lambda\phi_{13})z_x+\lambda_x\phi_1,$$

$$F_{xy} = f_{12}+\lambda\phi_{12}+(f_{13}+\lambda\phi_{13})z_y+\lambda_y\phi_1,$$

and

$$F_{xy} = f_{12}+\lambda\phi_{12}+(f_{23}+\lambda\phi_{23})z_x+\lambda_x\phi_2,$$
$$F_{yy} = f_{22}+\lambda\phi_{22}+(f_{23}+\lambda\phi_{23})z_y+\lambda_y\phi_2.$$

Differentiation of (19) similarly gives

$$0 = f_{13}+\lambda\phi_{13}+(f_{33}+\lambda\phi_{33})z_x+\lambda_x\phi_3,$$
$$0 = f_{23}+\lambda\phi_{23}+(f_{33}+\lambda\phi_{33})z_y+\lambda_y\phi_3.$$

In virtue of these six equations, if, in the restricted Hessian

$$H(f;\phi) = \begin{vmatrix} f_{11}+\lambda\phi_{11} & f_{12}+\lambda\phi_{12} & f_{13}+\lambda\phi_{13} & \phi_1 \\ f_{12}+\lambda\phi_{12} & f_{22}+\lambda\phi_{22} & f_{23}+\lambda\phi_{23} & \phi_2 \\ f_{13}+\lambda\phi_{13} & f_{23}+\lambda\phi_{23} & f_{33}+\lambda\phi_{33} & \phi_3 \\ \phi_1 & \phi_2 & \phi_3 & 0 \end{vmatrix},$$

we multiply the last two columns by z_x, λ_x respectively and add to the first column, and then multiply the same two columns by z_y, λ_y respectively and add to the second column, we get

$$H(f;\phi) = \begin{vmatrix} F_{xx} & F_{xy} & f_{13}+\lambda\phi_{13} & \phi_1 \\ F_{xy} & F_{yy} & f_{23}+\lambda\phi_{23} & \phi_2 \\ 0 & 0 & f_{33}+\lambda\phi_{33} & \phi_3 \\ 0 & 0 & \phi_3 & 0 \end{vmatrix} = -\phi_3^2 \begin{vmatrix} F_{xx} & F_{xy} \\ F_{xy} & F_{yy} \end{vmatrix}.$$

In like fashion that principal minor of $H(f;\phi)$ obtained by omitting the second row and column reduces to

$$-\phi_3^2 F_{xx}.$$

For future reference we enunciate in explicit terms the theorem we have now proved regarding the Hessian of an implicit function:

(20) *If $f(x,y,z) = F(x,y)$ when z is defined by $\phi(x,y,z) = 0$, then the Hessian of F is given by*

$$\begin{vmatrix} F_{xx} & F_{xy} \\ F_{xy} & F_{yy} \end{vmatrix} = -\phi_z^{-2} \begin{vmatrix} f_{xx}+\lambda\phi_{xx} & f_{xy}+\lambda\phi_{xy} & f_{xz}+\lambda\phi_{xz} & \phi_x \\ f_{xy}+\lambda\phi_{xy} & f_{yy}+\lambda\phi_{yy} & f_{yz}+\lambda\phi_{yz} & \phi_y \\ f_{xz}+\lambda\phi_{xz} & f_{yz}+\lambda\phi_{yz} & f_{zz}+\lambda\phi_{zz} & \phi_z \\ \phi_x & \phi_y & \phi_z & 0 \end{vmatrix},$$

where $\lambda = -f_z/\phi_z$.

The foregoing analysis breaks down, if ϕ_3 vanishes at the stationary point. We then choose some other variable, x say, for elimination between f and $\phi = 0$. There still remains the exceptional case in which ϕ_1, ϕ_2, ϕ_3 all vanish at the stationary point, but, as we shall see in the

following section, a special interpretation has to be put on this exceptional case.

In every other case the preceding analysis establishes the equivalence of the restricted Hessian and its principal minors to the Hessian and principal minors of the Hessian of the corresponding unrestricted function obtained by suitable elimination of variables. Hence, in the terms of chapter IX (27), (34), we may state conditions *sufficient* for a minimum:

(21) *A sufficient condition that a function $f(x_1,\ldots,x_n)$, restricted by m ($< n$) conditions $\phi_r(x_1,\ldots,x_n) = 0$, be a minimum at a stationary point is that the restricted Hessian and a sequence of principal minors have the sign $(-)^m$ at that point,*

and again conditions *necessary* for a minimum:

(22) *If a stationary point of $f(x_1,\ldots,x_n)$, restricted by m ($< n$) conditions $\phi_r(x_1,\ldots,x_n) = 0$, is a minimum, it is necessary that the restricted Hessian and every principal minor of every order either have the sign $(-)^m$ or be zero. Moreover, if any one of these determinants is not zero, the same must be true of its every principal minor.*

If the restricted Hessian has the sign $(-)^m$, it is then both necessary and sufficient for a minimum that every principal minor have this sign. Thus a gap is left between the conditions (21), (22) only in the case of the restricted Hessian's vanishing. It is this case which we reserve for subsequent discussion as the semi-definite case of a restricted minimum.

4. Inversion of the conditions

The foregoing apparatus for determining minima, in so far as it rests on equations (6) and the restricted Hessian (15), depends only on derivatives of f, ϕ and is therefore unaltered in external form, if we replace the equation of condition $\phi = 0$ by the rather more general equation $\phi = b$, where b is a constant. Actually, of course, the equation of condition has to be associated with equations (7) to determine the coordinates of the stationary point, and therefore variation of the equation of condition varies the coordinates of the stationary point and therefore also the values, taken at the stationary point, of the restricted Hessian and its principal minors.

Now suppose that we interchange the functions f, ϕ and seek the minima of ϕ subject to condition $f =$ constant. This interchange has no effect on the set of equations (7)

$$\left\|\begin{matrix} f_1 & . & . & . & f_n \\ \phi_1 & . & . & . & \phi_n \end{matrix}\right\| = 0,$$

which with the equation $f =$ constant, determine the coordinates of the stationary point. Hence, if the constants be properly chosen, a stationary point of f when ϕ is constant is also a stationary point of ϕ when f is constant.

In general and more precisely, we may say that

(23) *If $(\xi_1,\ldots,\xi_n)$ is a stationary point of $f(x_1,\ldots,x_n)$ subject to m $(< n)$ conditions*
$$\phi_r(x_1,\ldots,x_n) = a_r \qquad (r = 1,\ldots,m),$$
it is also a stationary point of $\phi_s(x_1,\ldots,x_n)$ subject to the m conditions
$$f(x_1,\ldots,x_n) = f(\xi_1,\ldots,\xi_n),$$
$$\phi_r(x_1,\ldots,x_n) = a_r \qquad (r = 1,\ldots,s-1,s+1,\ldots,m),$$
and the stationary value $\phi_s(\xi_1,\ldots,\xi_n)$ is a_s.

We return to the simplified case of $m = 1$ and examine the stationary point for a minimum of ϕ. The restricted Hessian of ϕ is
$$H_\phi = \begin{vmatrix} \phi_{11}+\lambda' f_{11} & \cdot & f_1 \\ \cdot & \cdot & \cdot \\ f_1 & \cdot & 0 \end{vmatrix},$$
where λ' is defined by the n equations
$$\phi_r+\lambda' f_r = 0 \qquad (r = 1,\ldots,n), \tag{24}$$
which are consistent if evaluated at a stationary point of ϕ. If we compare these n equations with the corresponding equations (6) for f, namely
$$f_r+\lambda\phi_r = 0 \qquad (r = 1,\ldots,n),$$
it is evident that the two sets of equations, *evaluated at the same stationary point*, give
$$\lambda' = 1/\lambda.$$
Substituting for λ', f_r in H_ϕ we get
$$H_\phi = \lambda^2 \begin{vmatrix} (f_{11}+\lambda\phi_{11})/\lambda & \cdot & \phi_1 \\ \cdot & \cdot & \cdot \\ \phi_1 & \cdot & 0 \end{vmatrix},$$
which is effectively the restricted Hessian of f/λ_0, where λ_0 is λ evaluated at the stationary point and therefore to be regarded as a constant, and the positive factor λ_0^2 can be ignored.

Hence, if we exclude the semi-definite case of vanishing Hessian, conditions necessary to ensure a minimum ϕ at the common stationary point are also sufficient to ensure a minimum of $\lambda_0 f$ there and conversely. Thus, if λ_0 is positive, maxima and minima of f correspond directly to maxima and minima of ϕ; if λ_0 is negative, maxima and minima of f correspond, criss-cross, to minima and maxima of ϕ.

Incorporating (23) we may state the result for the general case in the form:

(25) *If* $(\xi_1,\ldots,\xi_n)$ *is a minimum of* $f(x_1,\ldots,x_n)$ *subject to the* m $(< n)$ *conditions*

$$\phi_r(x_1,\ldots,x_n) = a_r \qquad (r = 1,\ldots,m),$$

it is also a minimum or maximum of $\phi_s(x_1,\ldots,x_n)$ *subject to the* m *conditions*

$$f(x_1,\ldots,x_n) = f(\xi_1,\ldots,\xi_n),$$

$$\phi_r(x_1,\ldots,x_n) = a_r \qquad (r = 1,\ldots,s-1,s+1,\ldots,m),$$

according as λ_s *defined by the* n *equations*

$$\frac{\partial f}{\partial x_p}+\sum_{r=1}^{s-1}\lambda_r\frac{\partial\phi_r}{\partial x_p}+\lambda_s\frac{\partial\phi_s}{\partial x_p}+\sum_{r=s+1}^{m}\lambda_r\frac{\partial\phi_r}{\partial x_p}=0 \qquad (p = 1,\ldots,n)$$

is positive or negative. Moreover, the minimum or maximum value of ϕ_s *is* a_s.

[*The semi-definite case is supposed excluded.*]

5. An exceptional case and an example

To return to our simplified case of $m = 1$, the theorem (25) gives no result, if λ is zero at the stationary point. Moreover, the corresponding λ' is apparently infinite: actually therefore equations (24) give no solution for λ'. In explanation of this anomaly we observe that, if $\lambda = 0$, we have from (6) the set of equations

$$f_1, f_2,\ldots,f_n = 0,$$

i.e. the stationary point in question is also a stationary point of the unrestricted f.

If now this stationary point $(\xi_1,\ldots,\xi_n)$ is actually a turning value of unrestricted f, then, by definition, the equation

$$f(x_1,\ldots,x_n) = f(\xi_1,\ldots,\xi_n) \tag{26}$$

is not satisfied in the neighbourhood of $(\xi_1,\ldots,\xi_n)$ but only at $(\xi_1,\ldots,\xi_n)$ itself. Hence the domain of any function $\phi(x_1,\ldots,x_n)$ subjected to these restrictions reduces to the solitary point $(\xi_1,\ldots,\xi_n)$ and questions about its maxima and minima there become unintelligible.

If, however, the stationary point is not actually a turning-point of unrestricted f, we must seek a further explanation of the phenomenon. Consider most simply a function $f(x,y)$ of two variables subject to a single restriction $\phi(x,y) =$ constant. At a stationary point (ξ,η) of the unrestricted function $f(x,y)$ we have $f_x = 0 = f_y$ and therefore, as we have seen, the point is a double point of the curve $f(x,y) = f(\xi,\eta)$ and has real branches, if the point is not a turning-point of $f(x,y)$.

Suppose most simply that this double point is a real node. The two branches have directions given by

$$f_{xx}(Dx)^2+2f_{xy}D_xD_y+f_{yy}(Dy)^2=0,$$

say $$\alpha\,Dx+\beta\,Dy=0, \qquad \alpha'\,Dx+\beta'\,Dy=0.$$

It is then *one* of these two last equations and not the evanescent equation

$$f_x\,Dx+f_y\,Dy=0$$

that we ought to combine with $\phi_x\,Dx+\phi_y\,Dy=0$ to determine λ. There are consequently two possible values of λ and two restricted Hessians to be considered. Clearly ϕ is not a minimum at the point, unless it is similarly a minimum considered along *both* branches of $f=$ constant At a multiple point of higher order there will, of course, be correspondingly more possible values of λ and more numerous discriminations of maxima and minima to be made. Similar though less simple conditions apply to functions of more variables.

As a simple illustration of the theory of restricted minima, consider the three functions

$$2f,2\phi,2\psi=(x-a_r)^2+(y-b_r)^2+(z-c_r)^2 \qquad (r=1,2,3).$$

At a stationary value of f restricted by $\phi=$ constant, $\psi=$ constant we have

$$x-a_1+\lambda(x-a_2)+\mu(x-a_3)=0,\ \text{etc.}$$

A stationary point is therefore given by

$$\left(\frac{a_1+\lambda a_2+\mu a_3}{1+\lambda+\mu},\ \frac{b_1+\lambda b_2+\mu b_3}{1+\lambda+\mu},\ \frac{c_1+\lambda c_2+\mu c_3}{1+\lambda+\mu}\right),$$

i.e. is the point P whose areal coordinates are $1/(1+\lambda+\mu)$, $\lambda/(1+\lambda+\mu)$, $\mu/(1+\lambda+\mu)$ referred to the triangle $A_1A_2A_3$ whose vertices are (a_r,b_r,c_r) $(r=1,2,3)$.

The restricted Hessian reduces to

$$(1+\lambda+\mu)\sum_{x,y,z}\begin{vmatrix} y-b_2 & y-b_3 \\ z-c_2 & z-c_3 \end{vmatrix}^2$$

and so has the sign of $1+\lambda+\mu$. Its first minors are

$$\begin{vmatrix} y-b_2 & y-b_3 \\ z-c_2 & z-c_3 \end{vmatrix}^2, \quad \text{etc.}$$

and are therefore positive.

Now $1+\lambda+\mu$ is positive or negative according as P and A_1 are on the same or opposite sides of the line A_2A_3. Hence P is a minimum or a maximum of restricted f according as P and A_1 are on the same or opposite sides of A_2A_3.

Again λ is positive or negative according as $A_3 P$ (produced if necessary) lies in the internal or external angle A_3 of the triangle $A_1 A_2 A_3$. Hence, if P lies in the internal angle or the internal opposite angle at A_3, maxima and minima of f when ϕ, ψ are constant correspond to maxima and minima of ϕ when f, ψ are constant. If P be in either of the external angles at A_3, the maxima and minima of f, ϕ correspond crosswise.

If $\lambda = 0$, P lies on $A_1 A_3$. The spheres $f =$ constant, $\phi =$ constant touch at P and therefore have as their real curve of intersection the solitary point P. In this case, then, the question of maxima and minima of ϕ subject to the restrictions $f =$ constant, $\psi =$ constant does not arise.

These results are at once confirmed from a figure by elementary geometrical considerations.

6. The semi-definite case. Functions of two variables

We come now to the long-postponed 'semi-definite' or 'cuspidal' case, beginning with $f(x, y)$ a function of two variables. The position, as we left it in chapter IX, was as follows. The function $f(x, y)$ is being examined at a stationary point. At this point $f_x = 0 = f_y$, and therefore, along any curve through the point,

$$Df = 0 \quad \text{and} \quad D^2 f = f_{xx}(Dx)^2 + 2f_{xy}\,DxDy + f_{yy}(Dy)^2.$$

Moreover, $f_{xx}, f_{yy} \geqslant 0$, $f_{xx}f_{yy} - f_{xy}^2 = 0$, so that we may write

$$D^2 f = \{\sqrt{(f_{xx})}Dx + \sqrt{(f_{yy})}Dy\}^2.$$

Thus $D^2 f = 0$ along curves in the direction given by

$$\sqrt{(f_{xx})}Dx + \sqrt{(f_{yy})}Dy = 0, \tag{27}$$

but $D^2 f > 0$ along every other curve. The point is therefore a minimum of the function considered along any curve not in the direction (27). Along any curve in this direction the character of the point is as yet indeterminate and we need to examine derivatives of higher order.

The practical problem is definite enough: we have to determine whether in the last analysis the stationary point is or is not an isolated point of the curve $f(x, y) =$ constant; this practical problem we can often solve by various devices of approximation. To the theoretical problem we can also give an exact answer *in general*, i.e. provided that f_{xx}, f_{xy}, f_{yy} are not *all* zero at the point.†

Let us recall that for a function f of one variable the exact condition for a minimum at a stationary point is that the first of the set of derivatives

$$D^2 f,\ D^3 f,\ \ldots,\ D^n f$$

that does not vanish at the point be of even order and be positive.

† The exceptional case in which *every* second derivative vanishes at the stationary point is briefly considered below in § 13.

For the function of two variables I shall similarly define in general a set of expressions
$$H_2, H_3, \ldots, H_n$$
(of which H_2 will be the Hessian itself), such that a stationary point will be a minimum, if and only if the first H_r which does not vanish at the point is of even order and is positive.

I shall appeal continually to the theorem of chapter VI (45) as a lemma, writing
$$D_r \equiv D^r/r!, \qquad x_r \equiv D_r x, \qquad y_r \equiv D_r y,$$
and expressing $D_n f$ as a function of variables $x, y, x_1, y_1, \ldots, x_n, y_n$ regarded as all independent. The lemma then states that, if $r < n$,
$$\left(\frac{\partial}{\partial x_r}, \frac{\partial}{\partial y_r}\right) D_n f = D_{n-r}(f_x, f_y). \tag{28}$$
For partial derivatives of the higher orders I shall now use the more convenient notation
$$f_{pq} = \frac{\partial^{p+q} f}{\partial x^p \partial y^q}.$$

7. A simpler special case. Definition of $H_n(x, y)$

Let us begin by considering the simpler special case in which at the stationary point
$$f_{20} = 1, \qquad f_{11} = 0, \qquad f_{02} = 0. \tag{29}$$
The *critical direction* in which alone the semi-definite quadratic
$$D^2 f = f_{20}\,(Dx)^2 + 2f_{11}\,DxDy + f_{02}\,(Dy)^2$$
actually vanishes is now $Dx = 0$. In view of (29) we may define it as the direction $Df_x = 0$.

More generally, we shall require to examine the behaviour of f along *critical curves* whose derivatives satisfy the conditions
$$D_s f_x = 0 \qquad (s = 1, \ldots, r). \tag{30}$$
The highest derivatives involved are x_r, y_r. By the lemma we have
$$\frac{\partial}{\partial x_r}(D_r f_x) = f_{20} = 1, \qquad \frac{\partial}{\partial y_r}(D_r f_x) = f_{11} = 0,$$
so that $D_r f_x = 0$ may be written in the form
$$x_r = A_r(x_{r-1}, \ldots, x_1; y_{r-1}, \ldots, y_1), \tag{31}$$
and similarly any $D_s f_x = 0$ may be written as
$$x_s = A_s(x_{s-1}, \ldots, x_1; y_{s-1}, \ldots, y_1). \tag{32}$$
Accordingly we may use equations (30) to substitute in succession for $x_r, x_{r-1}, \ldots, x_1$, leaving at length only the variables $y_{r-1}, \ldots, y_1$. I shall denote by $[F]_r$ the result of such a substitution on an expression F. The

substitution itself I shall call X_r, and a curve whose derivatives satisfy the relations (30) I shall call a curve K_r.

Now the lemma at once gives

$$\frac{\partial}{\partial y_q}(D_p f_x) = \frac{\partial}{\partial y_{q-m}}(D_{p-m} f_x) \qquad (p \geqslant q \geqslant m),$$

and our first step is to see that this relation persists when we apply the respective substitutions X_{p-1}, X_{p-m-1} to the two sides, i.e. to show that

$$\frac{\partial}{\partial y_q}[D_p f_x]_{p-1} = \frac{\partial}{\partial y_{q-m}}[D_{p-m} f_x]_{p-m-1} \qquad (p \geqslant q \geqslant m). \tag{33}$$

We must remember that y_q enters $[F]_r$ both directly and also through every substitution

$$x_s = A_s(x_{s-1},\ldots,x_1; y_s,\ldots,y_1) \qquad (r \geqslant s \geqslant q),$$

or, more precisely, through $\quad x_s = [A_s]_{s-1}$,
since we may suppose $x_1, x_2,\ldots, x_r$ replaced, successively and in this order, by their equivalent expressions in $y_1,\ldots, y_r$. Thus

$$\frac{\partial}{\partial y_q}[D_p f_x]_{p-1} = \left[\frac{\partial}{\partial y_q}(D_p f_x)\right]_{p-1} + \sum_{s \geqslant q}^{s \leqslant p-1}\left\{\left[\frac{\partial}{\partial x_s}(D_p f_x)\right]_{p-1}\frac{\partial}{\partial y_q}[A_s]_{s-1}\right\}. \tag{34}$$

The summation $\sum$ ends above at $s = p-1$, since the substitution X_{p-1} involves x's up to x_{p-1} only; the summation ends below at $s = q$, since the substitutions X_{q-1}, X_{q-2},... introduce y's of suffixes less than q and are therefore not concerned in the partial differentiations with respect to y_q.

Firstly, if $p = q$, we obtain our result (33) at once, for in (34) there are now no terms in $\sum$, since the conditions $p-1 \geqslant s \geqslant p$ are incompatible, and we have merely

$$\begin{aligned}\frac{\partial}{\partial y_p}[D_p f_x]_{p-1} &= \left[\frac{\partial}{\partial y_p}(D_p f_x)\right]_{p-1} \\ &= \left[\frac{\partial}{\partial y_{p-m}}(D_{p-m} f_x)\right]_{p-1}, \text{ by the lemma,} \\ &= \left[\frac{\partial}{\partial y_{p-m}}(D_{p-m} f_x)\right]_{p-m-1},\end{aligned}$$

since $p-m-1$ is now the highest suffix of x. Thus

$$\frac{\partial}{\partial y_p}[D_p f_x]_{p-1} = \frac{\partial}{\partial y_{p-m}}[D_{p-m} f_x]_{p-m-1}. \tag{35}$$

If $p > q$, I use an inductive argument: fixing q, I show that (33) is true for p, if it is true for $p-1$, $p-2$,.... Now by definition

$$A_s \equiv x_s - D_s f_x,$$

and so, for any term in the summation of (34),

$$\frac{\partial}{\partial y_q}[A_s]_{s-1} = -\frac{\partial}{\partial y_q}[D_s f_x]_{s-1} = -\frac{\partial}{\partial y_{q-m}}[D_{s-m} f_x]_{s-m-1} \tag{36}$$

by the hypothesis of the induction, since $s \leqslant p-1$ in the summation of (34). Thus, by definition of A_{s-m},

$$\frac{\partial}{\partial y_q}[A_s]_{s-1} = \frac{\partial}{\partial y_{q-m}}[A_{s-m}]_{s-m-1}.$$

Again, as in the proof of (35),

$$\left[\frac{\partial}{\partial y_q}(D_p f_x)\right]_{p-1} = \left[\frac{\partial}{\partial y_{q-m}}(D_{p-m} f_x)\right]_{p-m-1}$$

and
$$\left[\frac{\partial}{\partial x_s}(D_p f_x)\right]_{p-1} = \left[\frac{\partial}{\partial x_{s-m}}(D_{p-m} f_x)\right]_{p-m-1}. \tag{37}$$

Thus, by (34), (36), (37),

$$\begin{aligned}\frac{\partial}{\partial y_q}[D_p f_x]_{p-1} &= \left[\frac{\partial}{\partial y_{q-m}}(D_{p-m} f_x)\right]_{p-m-1} + \\ &\quad + \sum_{s-m \geqslant q-m}^{s-m \leqslant p-m-1} \left[\frac{\partial}{\partial x_{s-m}}(D_{p-m} f_x)\right]_{p-m-1} \frac{\partial}{\partial y_{q-m}}[A_{s-m}]_{s-m-1} \\ &= \frac{\partial}{\partial y_{q-m}}[D_{p-m} f_x]_{p-m-1}, \quad \text{by comparison of (34).}\end{aligned}$$

Thus the induction is established, and, since (33) is true for $p = q$, it is proved for any p not less than q.

In similar fashion

$$\frac{\partial}{\partial y_q}[D_p f]_p = \left[\frac{\partial}{\partial y_q}(D_p f)\right]_p + \sum_{s \geqslant q}^{s \leqslant p} \left[\frac{\partial}{\partial x_s}(D_p f)\right]_p \frac{\partial}{\partial y_q}[A_s]_{s-1},$$

and it follows immediately that

$$\frac{\partial}{\partial y_q}[D_p f]_p = \frac{\partial}{\partial y_{q-m}}[D_{p-m} f]_{p-m} \qquad (p \geqslant q \geqslant m). \tag{38}$$

We now examine the hypothesis that simultaneously

$$\frac{\partial}{\partial y_1}[D_s f]_s = 0 \qquad (s = 1,..., p-1). \tag{39}$$

Then, by (38),

$$\frac{\partial}{\partial y_r}[D_p f]_p = \frac{\partial}{\partial y_1}[D_{p-r+1}f]_{p-r+1}$$
$$= 0, \quad \text{by the hypothesis, if } p \geqslant r > 1.$$

Thus $[D_p f]_p$ is a function of y_1 alone. Dimensional considerations show that it is a pure power of y_1. More precisely we write

$$[D_p f]_p = y_1^p H_p(x_0, y_0), \tag{40}$$

where H_p involves only partial derivatives of f evaluated at the stationary point (x_0, y_0). Hence, if $H_p(x_0, y_0) = 0$, we shall have

$$[D_p f]_p = 0, \qquad \frac{\partial}{\partial y_1}[D_p f]_p = 0,$$

and we can extend the hypothesis (39) up to $s = p$.

More generally, if we have the set of conditions

$$H_r(x_0, y_0) = 0 \quad (r = 1,\dots, 2n-1), \qquad H_{2n}(x_0, y_0) > 0, \tag{41}$$

we can deduce inductively that

$$[D_r f]_r = 0 \quad (r = 1,\dots, 2n-1), \qquad [D_{2n} f]_{2n} > 0. \tag{42}$$

8. Sufficiency of the conditions

I now show that the conditions (41) are *sufficient* to secure a minimum of $f(x, y)$ at the stationary point (x_0, y_0).

In the first place, by considering dimensions, we see that $D_{2r+1}f$, $D_{2r}f$ are both linear in $x_{r+1},\dots, x_{2r}$. Moreover, by the lemma,

$$\frac{\partial}{\partial x_{r+s}}(D_{2r+1}f) = D_{r-s+1}f_x \qquad (r > s \geqslant 1).$$

Thus the substitution X_r clears $D_{2r+1}f$ of $x_{r+1},\dots, x_{2r}$ as well as of $x_1,\dots, x_r$: that is to say,

$$[D_{2r+1}f]_{2r+1} = [D_{2r+1}f]_r.$$

Similarly,

$$[D_{2r}f]_{2r} = [D_{2r}f]_r.$$

Again $D_{2r}f$ is quadratic in x_r, and, by the lemma,

$$\frac{\partial}{\partial x_r}(D_{2r}f) = D_r f_x = x_r - A_r, \quad \text{by (31)}.$$

Thus

$$D_{2r}f = \tfrac{1}{2}(x_r - A_r)^2 + B_r,$$

where B_r, like A_r, involves $x_1,\dots, x_{r-1}$ but not $x_r,\dots, x_{2r}$, and so

$$[D_{2r}f]_{r-1} = \tfrac{1}{2}\{x_r - [A_r]_{r-1}\}^2 + [B_r]_{r-1}.$$

The further substitution $x_r = [A_r]_{r-1}$ gives

$$[D_{2r}f]_r = [B_r]_{r-1},$$

and so

$$[D_{2r}f]_{r-1} = [D_{2r}f]_r + \tfrac{1}{2}\{x_r - [A_r]_{r-1}\}^2. \tag{43}$$

Now the curves K_r are a sub-class of the curves K_{r-1}, and (43) shows that, along those curves K_{r-1} that are not also members of the sub-class K_r,
$$[D_{2r}f]_{r-1} > 0, \quad \text{if} \quad [D_{2r}f]_r \geqslant 0.$$
If $s \leqslant 2r-1$, $\qquad [D_s f]_{r-1} = [D_s f]_s = 0$, by (42).

Thus, along a curve K_{r-1} ($r < n$) that is not also a curve K_r, the earliest derivative of f that does not vanish at the stationary point is $[D_{2r}f]_{r-1}$, which is positive and of even order: along *all* curves K_n, the earliest derivative that does not vanish at the stationary point is $[D_{2n}f]_n$, which again is positive and of even order.

We have now accounted for all curves through the stationary point in which $D^{2n}x$, $D^{2n}y$ exist and Dx, Dy are not both zero. Along all these curves we have shown that $f(x,y)$, under the conditions (42), is a minimum at the stationary point. Thus, according to our general theory,† conditions (42) are sufficient to secure $f(x,y)$ a minimum at the stationary point

Conversely, if along some K_n
$$[D^r f]_n = 0 \qquad (r < 2n-1).$$
then, since f has to be a minimum along this curve, we must have $[D^{2n}f]_n > 0$. Hence necessarily for a minimum the earliest $H_r(x_0, y_0)$ that does not vanish must be positive and of even suffix.

What we have now proved may be summed up in the theorem:

(44) *The necessary and sufficient condition that $f(x,y)$ have a minimum at a stationary point is that the first term of the sequence $\{H_n(x,y)\}$ that does not vanish at the stationary point be of even order and be positive at the point.*

The theorem, of course, becomes meaningless, if *every* H_n vanishes at the stationary point. That is a possibility to which we recur later. We consider first how to calculate H_n

9. Explicit definition of H_n

We no longer now restrict ourselves to the stationary point (x_0, y_0), but consider the general functions $H_n(x,y)$. The substitution $x_r = A_r$ appropriate to the stationary point must accordingly be replaced by its complete form
$$D_r f_x = f_{20}x_r + f_{11}y_r - A_r = 0,$$
or, let us say,
$$D_r f_x \equiv f_{20}(x_r - C_r), \tag{45}$$
where
$$C_r = C_r(x_{r-1},\ldots,x_1; y_r,\ldots,y_1; x, y).$$

† Chapter IX, § 4.

The substitution X_r, in this complete form, is, we shall see, unaffected by differentiation. The differential operator itself

$$D \equiv x_1\,\partial_x + y_1\,\partial_y$$

admits, of course, the substitution X_1. For brevity let us denote the transformed operator by D_1, i.e.

$$D_1 \equiv C_1\,\partial_x + y_1\,\partial_y.$$

We then prove that, for any polynomial ϕ,

$$[D\phi(x_r, x_{r-1},...)]_{r+1} = D_1\{[\phi(x_r, x_{r-1},...)]_r\}. \tag{46}$$

If, as usual, ϕ_1, ϕ_2,... denote the partial derivatives of ϕ in its first, second,... arguments, we have for the left-hand side of (46)

$$[(r+1)x_{r+1}\phi_1 + rx_r\phi_2 + ...]_{r+1} = (r+1)[C_{r+1}]_r[\phi_1]_r + r[C_r]_{r-1}[\phi_2]_r + ...$$

The right-hand side is

$$D_1\phi\{[C_r]_{r-1}, [C_{r-1}]_{r-2},...\} = D_1\{[C_r]_{r-1}\}.[\phi_1]_r + D_1\{[C_{r-1}]_{r-2}\}.[\phi_2]_r + ...$$

We have therefore to prove that

$$(s+1)[C_{s+1}]_s = D_1\{[C_s]_{s-1}\} \qquad (s = r, r-1,...), \tag{47}$$

which we prove inductively.

Assuming then that (46) holds up to $r-1$, we have, in particular,

$$[DC_r(x_{r-1}, x_{r-2},...)]_r = D_1\{[C_r(x_{r-1}, x_{r-2},...)]_{r-1}\},$$

i.e. briefly

$$[DC_r]_r = D_1\{[C_r]_{r-1}\}. \tag{48}$$

Now, from (45),

$$\begin{aligned}(r+1)f_{20}(x_{r+1} - C_{r+1}) &= (r+1)D_{r+1}f_x \\ &= D(D_r f_x) \\ &= D\{f_{20}(x_r - C_r)\} \\ &= f_{20}\{(r+1)x_{r+1} - DC_r\} + (x_r - C_r)Df_{20}.\end{aligned}$$

The substitution X_r gives

$$(r+1)[C_{r+1}]_r = [DC_r]_r,$$

i.e., with (48),

$$(r+1)[C_{r+1}]_r = D_1\{[C_r]_{r-1}\}.$$

Hence, if (47) holds up to $s = r-1$, it holds for $s = r$ also, and the induction is established. Since (46) evidently holds for $r = 0$, then (47) holds for $s = 0$. This furnishes the starting-point for the induction and so (46) is proved universally.

Applying this result to $D_n f$, we have, in virtue of (40),

$$(n+1)y_1^{n+1}H_{n+1} = (n+1)[D_{n+1}f]_{n+1} = D_1\{[D_n f]_n\} = D_1(H_n\,y_1^n).$$

Now $$D(f_x) = f_{20}x_1+f_{11}y_1 = f_{20}(x_1-C_1),$$
by the definition of C_1. Thus for D_1 we have
$$D_1 = (y_1/f_{20})(f_{20}\partial_y-f_{11}\partial_x),$$
and so $$(n+1)H_{n+1}y_1^{n+1} = (y_1/f_{20})(f_{20}\partial_y-f_{11}\partial_x)(H_n y_1^n). \tag{49}$$

We are not interested in H_{n+1} unless H_n vanishes. The effect of the equation $H_n = 0$ and the definition of H_{n+1} by differentiation of H_n are therefore not affected, if we multiply H_n by any factor which has finite first derivatives at the stationary point, and is not zero or infinite there. It is accordingly sufficient to define the sequence H_n by the recurrence-formula $$H_{n+1} = (f_{20}\partial_y-f_{11}\partial_x)H_n$$
and so to arrive at the definition
$$H_n = (f_{20}\partial_y-f_{11}\partial_x)^n f, \tag{50}$$
where after differentiation we put $f_{20} = 1$, $f_{11} = 0$.

10. The more general case of $H(x,y) = 0$.

We have now completed the theory for the special case in which
$$f_{20} = 1, \qquad f_{11} = 0, \qquad f_{02} = 0$$
at the stationary point. To extend this theory to the general case in which at the stationary point
$$H = 0, \qquad f_{20} > 0, \qquad f_{02} > 0 \qquad \text{(for a minimum)},$$
we swing the axes by the linear transformation
$$x = pX+p'Y, \qquad y = qX+q'Y,$$
$$\partial_X = p\partial_x+q\partial_y, \qquad \partial_Y = p'\partial_x+q'\partial_y,$$
where p, p', q, q' are constants such that at the stationary point
$$f_{XX} = (p\partial_x+q\partial_y)^2 f = 1, \qquad f_{YY} = (p'\partial_x+q'\partial_y)^2 f = 0,$$
$$f_{XY} = (p\partial_x+q\partial_y)(p'\partial_x+q'\partial_y)f = 0$$
These equations give at the stationary point
$$f_{20}/q'^2 = -f_{11}/p'q' = f_{02}/p'^2 = 1/(pq'-p'q)^2, \tag{51}$$
and at any point
$$f_{XX}\partial_Y-f_{XY}\partial_X = (p'q-pq')\{p(f_{11}\partial_x-f_{20}\partial_y)+q(f_{02}\partial_x-f_{11}\partial_y)\}.$$
Thus, sufficiently, we may write
$$H_n = \{p(f_{11}\partial_x-f_{20}\partial_y)+q(f_{02}\partial_x-f_{11}\partial_y)\}^n f. \tag{52}$$

For purposes of subsequent generalization it is more suggestive to write the operator in (52) as

$$\begin{vmatrix} f_{20} & f_{11} & -q \\ f_{11} & f_{02} & p \\ \partial_x & \partial_y & 0 \end{vmatrix}, \tag{53}$$

with the convention, of course, that ∂_x, ∂_y do not operate on the other elements of the determinant. With this convention we write (52) itself in the form

$$H_n = \begin{vmatrix} f_{20} & f_{11} & -q \\ f_{11} & f_{02} & p \\ \partial_x & \partial_y & 0 \end{vmatrix}^n f. \tag{54}$$

Now equations (51) fix the ratio of p'/q', but the ratio p/q is left undefined, so long, of course, as $p'q-pq' \neq 0$. In geometrical language, one of the new (X, Y)-axes is fixed as the cuspidal tangent, the other may be any other line through the cusp. We may therefore regard p, q in (52) as 'arbitrary constants' and expand in the powers $p^s q^{n-s}$ $(s = 0,..., n)$. This gives us $n+1$ coefficients, which are all alternative forms for H_n: they could be proved equivalent by use of the equations $H_r = 0$ $(r < n)$. On the other hand, by giving p, q special pairs of values we obtain infinitely many varieties of H_n, but at most $n+1$ of these varieties are linearly distinct.

Again, we may remark that the satisfaction of the sequence of conditions for a *particular* p, q is sufficient for a minimum: the satisfaction of the conditions for *every* p, q $(p'q-pq' \neq 0)$ is *necessary* for a minimum. In particular, then, they are satisfied for *every* p, q, if they are satisfied for *any* p, q. We may also note that the expressions H_{2n} regarded as polynomials in p, q, must maintain an invariable sign, under the given conditions, for all real p, q: exceptionally, they all vanish when $pq'-p'q = 0$.

To work out explicit forms for the H_n we may use indifferently the operators

$$f_{11}\partial_x - f_{20}\partial_y, \qquad f_{02}\partial_x - f_{11}\partial_y,$$

as we have said. More symmetrically we could take

$$H_n = \{\surd(f_{02})\partial_x - \surd(f_{20})\partial_y\}^n f,$$

in which every H of odd order is irrational. In any case the explicit forms rapidly become unmanageable. I give below (without proof) expressions for H_3, H_4, H_5, H_6: we may ignore H_1, H_2, since H_2 is the Hessian itself and H_1 vanishes automatically at the stationary point.

For brevity, write at the stationary point

$$f_{20} = a^2, \qquad f_{11} = ab, \qquad f_{02} = b^2,$$

and

$$\delta \equiv b\partial_x - a\partial_x.$$

Then we may obtain

$$\begin{aligned}
H_3 &= \delta^3 f,\\
H_4 &= \delta^4 f + 3\{\delta f_{20}\,\delta f_{02} - (\delta f_{11})^2\},\\
H_5 &= \delta^5 f + 5(\delta^2 f_{20}\,\delta f_{02} - 2\delta^2 f_{11}\,\delta f_{11} + \delta^2 f_{02}\,\delta f_{20}),\\
H_6 &= \delta^6 f + \tfrac{15}{2}(\delta^3 f_{20}\,\delta f_{02} - 2\delta^3 f_{11}\,\delta f_{11} + \delta^3 f_{02}\,\delta f_{20}) +\\
&\quad + 10\{\delta^2 f_{20} f_{02} - (\delta^2 f_{11})^2\} +\\
&\quad + \tfrac{5}{2}\{(\delta f_{02})^2 f_{40} - 4(\delta f_{02}\,\delta f_{11}) f_{31} + [4(\delta f_{11})^2 + 2\delta f_{20}\,\delta f_{02}] f_{22} -\\
&\qquad\qquad - 4(\delta f_{11}\,\delta f_{20}) f_{13} + (\delta f_{20})^2 f_{04}\} +\\
&\quad + 5\begin{vmatrix} \delta^2 f_{20} & \delta^2 f_{11} & \delta^2 f_{02} \\ f_{21} & f_{12} & f_{03} \\ f_{30} & f_{21} & f_{12} \end{vmatrix}.
\end{aligned}$$

11. The semi-definite case for functions of many variables

To indicate how the foregoing theory may be applied to functions of more than two variables, we can now outline its extension to $f(x,y,z)$ a function of three variables. Now the Hessian of f is the discriminant of D^2f regarded as a homogeneous quadratic in Dx, Dy, Dz. But, if the discriminant of such a quadratic vanishes, we know that a suitable linear transformation will reduce the quadratic to a quadratic in fewer variables. Hence, if the discriminant of $f(x,y,z)$ vanish at a stationary point, we can use a linear transformation of the variables to reduce the second derivative at that point to the simpler form

$$D^2f = f_{xx}\,(Dx)^2 + 2f_{xy}\,DxDy + f_{yy}\,(Dy)^2. \tag{55}$$

If, in addition, each principal minor of the Hessian also vanishes at the point, we can choose the linear transformation so that the second derivative at the point takes the still simpler form

$$D^2f = f_{xx}\,(Dx)^2. \tag{56}$$

But in this case *every* second derivative of $f(x_0, y, z)$, regarded as a function of two variables y, z only, namely f_{yy}, f_{yz}, f_{zz}, vanishes at the stationary point. Now this is a case in the theory of two variables that we have already postponed for subsequent consideration. We therefore similarly exclude the corresponding possibility (56) from our

present discussion and consider only the case (55) in which *the principal minors of the Hessian do not all vanish at the stationary point.*

It will be recalled that, for the function of three variables, the set of conditions postponed from chapter IX § 11 as the 'semi-definite case' were actually

$$f_{xx} \geqslant 0, \qquad f_{xx}f_{yy}-f_{xy}^2 \geqslant 0, \qquad H = 0,$$

where a choice of '$>$' imposed a similar choice *everywhere to the left.* Thus, since we have excluded the possibility (56), the conditions to be considered are only

$$f_{xx} > 0, \qquad f_{xx}f_{yy}-f_{xy}^2 > 0.$$

Hence D^2f in (55) is positive everywhere except along the curves in which both

$$x_1 = 0 = y_1,$$

i.e. the critical curves are now defined by

$$Df_x = 0 = Df_y,$$

and more generally, if we pursue the analogy, by

$$D^rf_x = 0 = D^rf_y.$$

We then find that along these curves we can write

$$[D^nf]_r = H_n z_1^n,$$

where H_n involves only f and its partial derivatives. This defines, for the function of three variables, the sequence H_n with characteristic properties.

Now $$D = x_1 \partial_x + y_1 \partial_y + z_1 \partial_z,$$

where $$0 = x_1 f_{xx} + y_1 f_{xy} + z_1 f_{xz},$$

$$0 = x_1 f_{xy} + y_1 f_{yy} + z_1 f_{yz}.$$

Rejecting an inconvenient factor we may sufficiently write D in the form

$$\begin{vmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{xy} & f_{yy} & f_{yz} \\ \partial_x & \partial_y & \partial_z \end{vmatrix},$$

and the sequence H_n is therefore defined as

$$H_n = \begin{vmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{xy} & f_{yy} & f_{yz} \\ \partial_x & \partial_y & \partial_z \end{vmatrix}^n f. \qquad (57)$$

We have so far been considering the special case in which D^2f has the reduced form (55). To obtain the formula for the general case we

impose a general linear transformation on (57), which gives ultimately the formula

$$H_n = \begin{vmatrix} f_{xx} & f_{xy} & f_{xz} & p \\ f_{xy} & f_{yy} & f_{yz} & q \\ f_{xz} & f_{yz} & f_{zz} & r \\ \partial_x & \partial_y & \partial_z & 0 \end{vmatrix}^n f, \tag{58}$$

where p, q, r are arbitrary constants subject only to the condition that the operator in (58) do not vanish identically at the stationary point. Evidently (58) gives a natural generalization of (54).

12. The semi-definite case for an implicit function

We consider here only the simplest case of a restricted minimum, namely that in which we require the minima of $f(x,y,z)$ subject to the condition

$$\phi(x,y,z) = 0. \tag{59}$$

From the form and theory of the 'restricted' Hessian, as developed in § 3 above, we are led to consider, as the natural generalization of (54) to the case of a restricted minimum, the sequence $\Delta^n f(x,y,z)$, where, in the notation of § 3,

$$\Delta = \begin{vmatrix} f_{11}+\lambda\phi_{11} & f_{12}+\lambda\phi_{12} & f_{13}+\lambda\phi_{13} & \phi_1 & p \\ f_{12}+\lambda\phi_{12} & f_{22}+\lambda\phi_{22} & f_{23}+\lambda\phi_{23} & \phi_2 & q \\ f_{13}+\lambda\phi_{13} & f_{23}+\lambda\phi_{23} & f_{33}+\lambda\phi_{33} & \phi_3 & r \\ \phi_1 & \phi_2 & \phi_3 & 0 & 0 \\ \partial_x & \partial_y & \partial_z & 0 & 0 \end{vmatrix}, \tag{60}$$

p, q, r being arbitrary constants and λ being any of $-f_1/\phi_1$, $-f_2/\phi_2$, $-f_3/\phi_3$, which are all equal at the stationary point. We proceed to show that this form imagined for H_n is actually a correct one.

If we expand Δ by the last column, we may write

$$\Delta = p\Delta_1+q\Delta_2+r\Delta_3, \tag{61}$$

where, identically,

$$0 = \phi_1\Delta_1+\phi_2\Delta_2+\phi_3\Delta_3. \tag{62}$$

Now, if Δ_1, Δ_2 are possible operators for defining a sequence H_n, then, by a usual argument, so are any multiples of them, in particular $(\phi_1/\phi_3)\Delta_1$, $(\phi_2/\phi_3)\Delta_2$, if we are careful to avoid zeros of ϕ_3. The sum of these two latter operators, again, is also a possible operator, and therefore, in virtue of (62), Δ_3 is a possible form of Δ, if Δ_1, Δ_2 are independently possible forms of Δ. Accordingly, to show that Δ can have the

general form (61), it is enough to show that it can have the less general form

$$\Delta = p\Delta_1+q\Delta_2.$$

This is the form that we now consider, i.e. we put $r = 0$ in (60). Now, as in § 3, let us assume that we can use the 'restrictive' condition (59) to eliminate z from $f(x,y,z)$, getting

$$F(x,y) \equiv f(x,y,z).$$

We then apply to the determinant (60), with $r = 0$, the operations applied in § 3 above† to the restricted Hessian, that is to say, we multiply the third and the fourth column by z_x, λ_x respectively and add to the first column, and then multiply the same two columns by z_y, λ_y respectively and add to the second column. We get, as in § 3, in virtue of the equations established there, that

$$\Delta_{r=0} = \begin{vmatrix} F_{xx} & F_{xy} & f_{13}+\lambda\phi_{13} & \phi_1 & p \\ F_{xy} & F_{yy} & f_{23}+\lambda\phi_{23} & \phi_2 & q \\ 0 & 0 & f_{33}+\lambda\phi_{33} & \phi_3 & 0 \\ 0 & 0 & \phi_3 & 0 & 0 \\ \partial_x+z_x\partial_z & \partial_y+z_y\partial_z & \partial_z & 0 & 0 \end{vmatrix}$$

$$= -\phi_3^2 \begin{vmatrix} F_{xx} & F_{xy} & p \\ F_{xy} & F_{yy} & q \\ \partial_x+z_x\partial_z & \partial_y+z_y\partial_z & 0 \end{vmatrix}.$$

But, for operation on $F(x,y)$ from which z has been eliminated, we replace the operators $\partial_x+z_x\partial_z$, $\partial_y+z_y\partial_z$ by the simple operators ∂_x, ∂_y; we may at the same time drop the factor $-\phi_3^2$ by our usual principle. Thus, sufficiently, we write

$$(\Delta_{r=0})^n f(x,y,z) = \begin{vmatrix} F_{xx} & F_{xy} & p \\ F_{xy} & F_{yy} & q \\ \partial_x & \partial_y & 0 \end{vmatrix}^n F(x,y).$$

The expression on the right exactly corresponds to the H_n defined in (54) above for determination of minima of the unrestricted $F(x,y)$. Thus the sequence $(\Delta_{r=0})^n f$ is a suitable sequence for determination of minima of $f(x,y,z)$ restricted by the condition $\phi(x,y,z) = 0$. As we have shown, if the sequence $(\Delta_{r=0})^n f$ is a suitable sequence, so more generally is $\Delta^n f$ for any constant r.

† p. 261.

We enunciate this formally in the theorem:

(63) *If the variables x, y, z are subject to the condition $\phi(x,y,z) = 0$, then the necessary and sufficient condition that a stationary value of $f(x,y,z)$ be a minimum is that, if λ denotes an arbitrary one of $-f_x/\phi_x$, $-f_y/\phi_y$, $-f_z/\phi_z$ and p, q, r are arbitrary constants, the first term of the sequence*

$$\begin{vmatrix} f_{xx}+\lambda\phi_{xx} & f_{xy}+\lambda\phi_{xy} & f_{xz}+\lambda\phi_{xz} & \phi_x & p \\ f_{xy}+\lambda\phi_{xy} & f_{yy}+\lambda\phi_{yy} & f_{yz}+\lambda\phi_{yz} & \phi_y & q \\ f_{xz}+\lambda\phi_{xz} & f_{yz}+\lambda\phi_{yz} & f_{zz}+\lambda\phi_{zz} & \phi_z & r \\ \phi_x & \phi_y & \phi_z & 0 & 0 \\ \partial_x & \partial_y & \partial_z & 0 & 0 \end{vmatrix}^n f(x,y,z)$$

that does not vanish at the stationary point is positive and of even index.

13. The residual cases

We have still to consider, however briefly, a class of cases left over from the foregoing discussions. These have purposely been left to the last, because they present problems of algebra that, for the most part, are still without satisfactory solution. They are typified by the case in which, for the function of two variables $f(x,y)$,

$$f_{xx}, \quad f_{xy}, \quad f_{yy}$$

all vanish at the stationary point: that is to say, in which D^2f vanishes along every curve at the stationary point. Then, for a turning value, we require, by the elementary theory of Chapter IX § 2, that D^3f also vanish along every curve at the stationary point, i.e. that

$$\left(\frac{\partial^3 f}{\partial x^3}, \frac{\partial^3 f}{\partial x^2 \partial y}, \frac{\partial^3 f}{\partial x \partial y^2}, \frac{\partial^3 f}{\partial y^3}\right)(Dx, Dy)^3 = 0 \tag{64}$$

for all (real) values of the ratio Dx/Dy. This is possible only if

$$\frac{\partial^3 f}{\partial x^3} = 0, \qquad \frac{\partial^3 f}{\partial x^2 \partial y} = 0, \qquad \frac{\partial^3 f}{\partial x \partial y^2} = 0, \qquad \frac{\partial^3 f}{\partial y^3} = 0 \tag{65}$$

simultaneously.

We then have at the stationary point

$$D^4 f = \left(\frac{\partial^4 f}{\partial x^4}, \frac{\partial^4 f}{\partial x^3 \partial y}, \frac{\partial^4 f}{\partial x^2 \partial y^2}, \frac{\partial^4 f}{\partial x \partial y^3}, \frac{\partial^4 f}{\partial y^4}\right)(Dx, Dy)^4, \tag{66}$$

and we require, for a turning value, that this quartic in Dx, Dy be of constant sign (zero excluded) for all (real) values of Dx, Dy except

simultaneous zeros, i.e. that $D^4f = 0$, regarded as a quartic equation in Dx/Dy, have no real roots.

Now the conditions that the roots of the quartic equation

$$a_0x^4+4a_1x^3+6a_2x^2+4a_3x+a_4 = 0 \tag{67}$$

be all imaginary may be given in the form†

$$\left.\begin{array}{lll} \Delta > 0, & H \geqslant 0; & \\ & H < 0, & a_0^2 I-12H^2 > 0; \\ \Delta = 0, & H > 0, & a_0^2 I-12H^2 = 0, \quad G = 0; \end{array}\right\} \tag{68}$$

where Δ is the discriminant and

$$G \equiv a_0^2a_3-3a_0a_1a_2+2a_1^3, \qquad H \equiv a_0a_2-a_1^2,$$
$$I \equiv a_0a_4-4a_1a_3+3a_2^2.$$

These conditions can be established by purely algebraic methods,‡ but it may not be inappropriate here to obtain them by the elementary machinery of the Calculus.

In the first place Δ, since it is proportional to the product of the squared differences of the roots of the equation, is positive or negative according as the number of pairs of conjugate imaginary roots is even or odd. Thus $\Delta > 0$ excludes the possibility of two real and two imaginary roots, and we only need some further condition that will distinguish between the cases of four real roots and four imaginary roots

For simplicity of working write $a_0x+a_1 \equiv x'$ and drop accents, so that the equation takes the 'reduced' form

$$x^4+6Hx^2+4Gx+a_0^2 I-3H^2 = 0. \tag{69}$$

Then the number of real roots of (69) will be the number of real intersections of the straight line

$$y+4Gx+a_0^2 I-3H^2 = 0 \tag{70}$$

with the quartic curve

$$y(x) = x^4+6Hx^2. \tag{71}$$

Now here

$$y'(x) = 4(x^3+3Hx), \qquad y''(x) = 12(x^2+H).$$

Hence, if $H \geqslant 0$, $y''(x)$ is never negative, $y'(x)$ is monotonic and so has only one zero, and therefore $y(x)$ has only one turning value. In this

† Cf. Burnside and Panton, *Theory of Equations*, 1 (1904), 144–5

‡ There is advantage in discussing such criteria geometrically by the method applied below to the quintic equation (Chapter XI worked example, p. 320): compare also Chaundy, *Quart. J. of Math.* (Oxford), 5 (1934), 10–22. That method is, in a sense, the geometric dual of the method used here.

case then the graph of $y(x)$ has the form of Fig. 1 and evidently any line such as (70) cuts it at most in two real points.

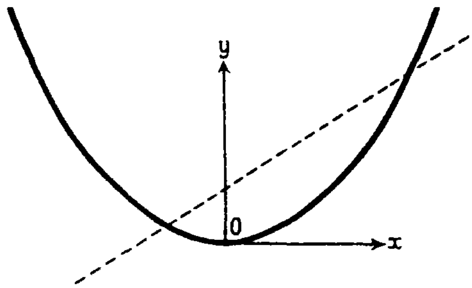

FIG. 1. ($H > 0$)

Thus $\Delta > 0$, $H > 0$ secures that all the roots of (69) shall be imaginary.

On the other hand, if $H < 0$, then $y(x)$ has turning values at $0, \pm\sqrt{(-3H)}$, the zeros of $y'(x)$. Of these, $x = 0$ gives a maximum, and $x = \pm\sqrt{(-3H)}$ give minima with a common tangent

$$y+9H^2 = 0.$$

Thus the graph of $y(x)$ is now as in Fig. 2.

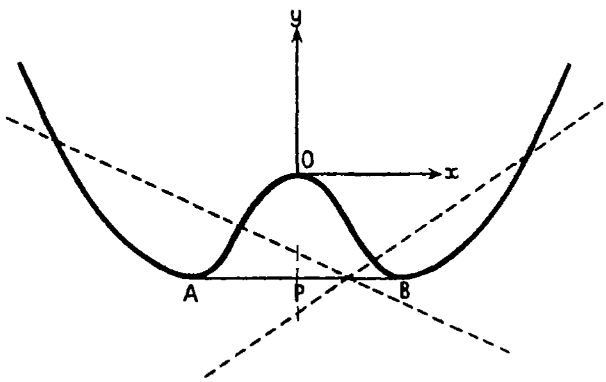

FIG. 2 ($H < 0$)

Now consider the point P where the double tangent AB cuts the y-axis. We see from the figure that any line cutting the y-axis above P† will cut the curve in *at least* two real points, while any line cutting the y-axis below P will cut the curve in *at most* two real points. Now the

† Or at P itself.

line (70) cuts the y-axis at $(0, -a_0^2 I + 3H^2)$ and P is the point $(0, -9H^2)$. Thus, if $\Delta > 0$, $H < 0$, the further condition

$$a_0^2 I > 12H^2$$

secures that all the roots of (69) shall be imaginary.

The remaining alternative of (68), namely

$$\Delta = 0, \qquad H > 0, \qquad a_0^2 I - 12H^2 = 0, \qquad G = 0$$

corresponds to the case of a double pair of conjugate imaginary roots in which (69) reduces to $(x^2+3H)^2 = 0.$

This completes the proof of (68).

It should be observed in the first place that although these conditions (68), are exact, i.e. necessary and sufficient, they are not unique: thus, we could have taken instead of P any point of the terminated line AB. Such a point still serves to separate lines cutting the curve *at least* twice from lines cutting it *at most* twice: P has merely been chosen as symmetrical and therefore most likely to give a convenient algebraic formula.

Again, the conditions (68) change their form for the reciprocal equation, i.e. the equation obtained by writing $1/x$ for x, and so, applied to the quartic (66), they give conditions asymmetrical in x, y.

Thus our algebraic problem, though fully solved, has a scarcely acceptable solution, and there is no advantage in stating it in its final form, by actually performing the substitution

$$(a_0, a_1, a_2, a_3, a_4) = \left(\frac{\partial^4 f}{\partial x^4}, \frac{\partial^4 f}{\partial x^3 \partial y}, \frac{\partial^4 f}{\partial x^2 \partial y^2}, \frac{\partial^4 f}{\partial x \partial y^3}, \frac{\partial^4 f}{\partial y^4}\right).$$

If now every *fourth* derivative vanishes at the stationary point, then as before, for a turning value, every fifth derivative must also vanish at the stationary point, and we are left with

$$D^6 f = \left(\frac{\partial^6 f}{\partial x^6}, \ldots, \frac{\partial^6 f}{\partial y^6}\right)(Dx, Dy)^6.$$

We have now to determine the conditions under which a sextic has an invariable sign. To this problem, although a method of approach is indicated by Sturm's theory of the distribution of the real roots of an algebraic equation, no proper answer has been found.

If we pass on to functions of more than two variables or to 'restricted' maxima and minima, new algebraic problems of like character arise: determination of the sign of a ternary quartic, of a 'restricted' quartic, etc. To none of these problems do I know the answer, and we must accordingly abandon the search for maxima and minima in the presence of these unsurmounted algebraic obstacles.

WORKED EXAMPLE

Show that the stationary values of the function $z(x, y)$ defined implicitly by the relation

$$A(z)x^2+2B(z)xy+C(z)y^2+k\{A(z)C(z)-B^2(z)\} = 0$$

are roots of the equation $A(z)C(z)-B^2(z) = 0.$

Show that a root z_0 of this equation such that $A(z_0)$ has the sign of k ($\neq 0$) and

$$A'(z_0)C'(z_0) > B'^2(z_0),$$

where dashes denote differentiation, is necessarily a stationary value of z and corresponds to a line of stationary points (x, y).

If $A(z)$, $B(z)$, $C(z)$ are differentiable to any order, show that every H_n vanishes at such stationary points, but that the stationary value is actually a turning value, being a maximum or a minimum according as $A(z_0)$, $A'(z_0)$ have the same or opposite signs.

Illustrate by considering the example in which

$$A(z) = z-a, \qquad B(z) = z-b, \qquad C(z) = z-c,$$

where a, b, c are constants.

Partial differentiation in x and in y gives

$$z_x = -\frac{2(Ax+By)}{\theta}, \qquad z_y = -\frac{2(Bx+Cy)}{\theta}, \tag{1}$$

where $\theta = A'x^2+2B'xy+C'y^2+k(A'C-2BB'+AC')$

and A, A',... are now written for $A(z)$, $A'(z)$,.... At a stationary point we must therefore have

$$Ax+By = 0, \qquad Bx+Cy = 0. \tag{2}$$

Substitution in the equation defining z shows that the corresponding stationary values z_0 are roots of the equation

$$A(z_0)C(z_0)-B^2(z_0) = 0. \tag{3}$$

If z_0 is a root of this equation such that $A(z_0)$ has the sign of k, we can write

$$A(z_0) = k\alpha^2, \qquad B(z_0) = k\alpha\beta, \qquad C(z_0) = k\beta^2,$$

where α, β are real and $\alpha \neq 0$. Thus

$$\theta(z_0) = A'x^2+2B'xy+C'y^2+k^2(A'\beta^2-2B'\alpha\beta+C'\alpha^2).$$

If, again, z_0 is such that $A'(z_0)C'(z_0) > B'^2(z_0)$,

then each quadratic in $\theta(z_0)$ has invariably the sign of $A'(z_0)$, unless both its arguments vanish. Since $\alpha \neq 0$, it follows that $\theta(z_0)$ is never zero and has the sign of $A'(z_0)$, and that consequently the conditions (2) are not only necessary but also sufficient for a stationary value.

Now (3) is the condition of consistency of (2) regarded as a pair of equations in x, y. Corresponding to such a stationary value z_0, there is thus a 'stationary edge' or line of stationary points

$$\frac{x}{y} = -\frac{B(z_0)}{A(z_0)} = -\frac{C(z_0)}{B(z_0)}. \tag{4}$$

For consideration of maxima and minima we need to know the higher derivatives, and we shall first show inductively that they have the general form given by

$$\frac{\partial^n z}{\partial x^p \partial y^{n-p}} = \frac{F\{x, y, A, B, C, \ldots, A^{(n)}, B^{(n)}, C^{(n)}\}}{\theta^{2n-1}}, \tag{5}$$

where $A^{(n)}$, $B^{(n)}$, $C^{(n)}$ are the nth derivatives in z of $A(z)$, $B(z)$, $C(z)$ and F is a polynomial function. Differentiation of (5) with respect to x gives

$$\frac{\partial^{n+1}z}{\partial x^{p+1}\partial y^{n-p}} = \frac{F_x+z_x[F_A A'+\ldots+F_C^{(n)}C^{(n+1)}]}{\theta^{2n-1}} -$$
$$-\frac{2n-1}{\theta^{2n}}F[2(A'x+B'y)+z_x\{A''x^2+2B''xy+C''y^2+$$
$$+k^2(A''C+2A'C'+AC''-2BB''-2B'^2)\}].$$

On substitution for z_x from (1) we see that

$$\frac{\partial^{n+1}z}{\partial x^{p+1}\partial y^{n-p}} = \frac{G\{x,y,A,B,C,\ldots,A^{(n+1)},B^{(n+1)},C^{(n+1)}\}}{\theta^{2n+1}},$$

where G is also some polynomial function. Thus the form of (5) is unaltered for partial differentiation in x (and similarly for partial differentiation in y). Comparison with (1) shows that (5) is true for $n = 1$: accordingly, by induction, it is true generally. Now the functions $A(z)$, $B(z)$, $C(z)$ are, by hypothesis, differentiable to any order and therefore also their derivatives are continuous and so bounded. Thus the numerator of (5) is bounded. But we have seen that $\theta(z_0)$ does not vanish for any stationary value z_0 that satisfies the conditions of the enunciation. Hence every derivative of z is bounded in the neighbourhood of any such z_0.

Now from (1) we have

$$z_{xx} = -\frac{2A}{\theta}-\frac{2(A'x+B'y)z_x}{\theta}+\frac{2(Ax+By)(\theta_x+\theta_z z_x)}{\theta^2},$$
$$z_{xy} = -\frac{2B}{\theta}-\frac{2(A'x+B'y)z_y}{\theta}+\frac{2(Ax+By)(\theta_y+\theta_z z_y)}{\theta^2},$$
$$z_{yy} = -\frac{2C}{\theta}-\frac{2(B'x+C'y)z_y}{\theta}+\frac{2(Bx+Cy)(\theta_y+\theta_z z_y)}{\theta^2},$$

which we may write

$$z_{xx} = -\frac{2A}{\theta}+\lambda_1 z_x,$$
$$z_{xy} = -\frac{2B}{\theta}+\lambda_2 z_y+\mu_1 z_x,$$
$$z_{yy} = -\frac{2C}{\theta}+\mu_2 z_y,$$

where λ_1, λ_2, μ_1, μ_2 are all of the general form

$$F\{x,y,A,B,C,A',B',C',..\}/\theta^n \qquad (6)$$

already considered in (5), so that they and their derivatives are bounded in the neighbourhood of those stationary values that are covered by the conditions of the enunciation. Hence

$$H_2 = \frac{4(AC-B^2)}{\theta^2}+\bar{\lambda}z_x+\bar{\mu}z_y, \quad \text{say},$$

where again $\bar{\lambda}$, $\bar{\mu}$ have the form (6). But from the equation defining z we have

$$k(AC-B^2) = -(Ax^2+2Bxy+Cy^2) = \tfrac{1}{2}\theta(xz_x+yz_y).$$

Thus we may write $\qquad H_2 = \lambda z_x+\mu z_y, \qquad (7)$

where λ, μ are also of the form (6).

Again, for the operator defining H_n write

$$\Delta \equiv p(z_{yy}\partial_x - z_{xy}\partial_y) + q(z_{xy}\partial_x - z_{xx}\partial_y),$$

where p, q are arbitrary constants. Then

$$\Delta z_x = pH_2, \qquad \Delta z_y = qH_2,$$

so that, from (7),

$$\begin{aligned} H_3 = \Delta H_2 &= z_x \Delta\lambda + z_y \Delta\mu + (p\lambda + q\mu)H_2 \\ &= (\Delta\lambda + p\lambda^2 + q\lambda\mu)z_x + (\Delta\mu + p\lambda\mu + q\mu^2)z_y \\ &= \lambda_3 z_x + \mu_3 z_y, \quad \text{say.} \end{aligned}$$

But $$\Delta\lambda = p(z_{yy}\lambda_x - z_{xy}\lambda_y) + q(z_{xy}\lambda_x - z_{xx}\lambda_y)$$

Since z_{xx}, z_{xy}, z_{yy}, λ are all of the form (6), so too is $\Delta\lambda$ and similarly $\Delta\mu$. Thus λ_3, μ_3 are also of the same form (6).

We may proceed in this way inductively, for if

$$H_n = \lambda_n z_x + \mu_n z_y,$$

then $$H_{n+1} = \lambda_{n+1} z_x + \mu_{n+1} z_y,$$

where $$\lambda_{n+1} = \Delta\lambda_n + \lambda(p\lambda_n + q\mu_n), \qquad \mu_{n+1} = \Delta\mu_n + \mu(p\lambda_n + q\mu_n).$$

Thus, by an inductive argument, every λ_n, μ_n is of the form (6) and is therefore finite at every stationary point that satisfies the conditions of the enunciation, and accordingly every H_n vanishes at every such stationary point.

To determine the actual behaviour of z in the neighbourhood of a stationary point z_0 we may expand the defining equation in powers of $z - z_0$ and write, to a first approximation,

$$\begin{aligned} &Ax^2 + 2Bxy + Cy^2 + k(AC - B^2) \\ &\quad + (z - z_0)\{A'x^2 + 2B'xy + C'y^2 + k(AC' - 2BB' + A'C)\} + O(z - z_0)^2 = 0, \end{aligned}$$

where A, A',.. are now evaluated at the stationary point. Thus $AC = B^2$ and

$$z - z_0 = -\frac{A^{-1}(Ax + By)^2 + O(z - z_0)^2}{\theta(z_0)}.$$

But, as we have already seen, $\theta(z_0)$ has the sign of $A'(z_0)$. Sufficiently near z_0, then, $z - z_0$ has the sign of $-A(z_0)A'(z_0)$, provided, of course, that we avoid the 'stationary edge' $A(z_0)x + B(z_0)y = 0$. In other words, such a stationary value z_0 is actually a turning value, and is a maximum or minimum according as $A(z_0)$, $A'(z_0)$ have the same or opposite signs.

Finally, in the given example in which

$$A(z) = z - a, \qquad B(z) = z - b, \qquad C(z) = z - c,$$

the defining relation is

$$(z - a)x^2 + 2(z - b)xy + (z - c)y^2 = k\{(a - 2b + c)z - (ac - b^2)\},$$

i.e. $$z = \frac{ax^2 + 2bxy + cy^2 - k(ac - b^2)}{(x + y)^2 - k(a - 2b + c)}.$$

There is a single stationary value

$$z_0 = \frac{ac - b^2}{a - 2b + c},$$

so that
$$z-z_0 = \frac{\{(a-b)x+(b-c)y\}^2}{(a-2b+c)\{(x+y)^2-k(a-2b+c)\}},$$
$$A(z_0) = -\frac{(a-b)^2}{a-2b+c}, \qquad A'(z_0) = 1,$$
$$A'(z_0)C'(z_0)-B'^2(z_0) = 0.$$

Hence, if $A(z_0)$ have the sign of k, i.e. if $k(a-2b+c)$ be negative, then $z-z_0$ has an invariable sign, namely that of $a-2b+c$, i.e. of $-A(z_0)$. In other words, z_0 is a maximum or minimum according as $A(z_0)$ is positive or negative, in conformity with the general theorem.

EXAMPLES X

1. If the elements of the determinant
$$\Delta = \begin{vmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{vmatrix}$$
obey the three relations
$$l_r^2+m_r^2+n_r^2 = 1 \qquad (r = 1,2,3),$$
show that the determinant is bounded and that its bounding values are ± 1. Interpret the result geometrically.

Discuss the bounding values of Δ when its constituents obey the single relation
$$\sum_{r=1}^{3}(l_r^2+m_r^2+n_r^2) = 3.$$

2. Obtain the bounding values of the determinant $|a_{rs}|$ of order n, when the constituents obey

(i) the n relations $\quad \sum_{s=1}^{n} a_{rs}^2 = c_r^2 \qquad (r = 1,\dots,n);$

(ii) the single relation $\quad \sum_{r=1}^{n}\sum_{s=1}^{n} a_{rs}^2 = nc^2.$

Show that the set of stationary values of the arguments a_{rs} can be determined with $\frac{1}{2}n(n-1)$ degrees of freedom in both (i) and (ii).

3. If the n variables $x_1,\dots,x_n$ obey any of the alternative sets of conditions

(i) $\sum_{r=1}^{n} a_r x_r = 1;$

(ii) $\sum_{r=1}^{n} a_r x_r = 0, \qquad \sum_{r=1}^{n} b_r x_r = 1;$

(iii) $\sum_{r=1}^{n} a_r x_r = 0, \qquad \sum_{r=1}^{n} b_r x_r = 0, \qquad \sum_{r=1}^{n} c_r x_r = 1;$

show that the function $\quad u \equiv x_1^2+\dots+x_n^2$

has a single stationary value, and that this value of the function is a minimum and least.

Obtain the stationary values in the forms

$$\text{(i)}\quad 1\Big/\Big\{\sum_r a_r^2\Big\},$$

$$\text{(ii)}\quad \Big\{\sum_r a_r^2\Big\}\Big/\Big\{\sum_{r,s}(a_r b_s - a_s b_r)^2\Big\},$$

$$\text{(iii)}\quad \left\{\sum_{r,s}\begin{vmatrix} a_r & b_r \\ a_s & b_s \end{vmatrix}^2\right\}\Bigg/\left\{\sum_{r,s,t}\begin{vmatrix} a_r & b_r & c_r \\ a_s & b_s & c_s \\ a_t & b_t & c_t \end{vmatrix}^2\right\}.$$

4. If the arguments $x_1,\dots,x_n$ obey either of the alternative sets of conditions

$$\text{(i)}\quad \sum_{r=1}^{n} x_n^2 = 1;$$

$$\text{(ii)}\quad \sum_{r=1}^{n} b_r x_r = 0,\qquad \sum_{r=1}^{n} x_n^2 = 1;$$

prove that the function $\qquad u = \sum_{r=1}^{n} a_r x_r$

has a pair of stationary values which are also the bounding values of the function, and show that these stationary values are given respectively by

$$\text{(i)}\quad u^2 = \sum a_r^2,\qquad \text{(ii)}\quad u^2\sum b_r^2 = \sum (a_r b_s - a_s b_r)^2.$$

5. If $\phi(x_1,\dots,x_n)$, $\psi(x_1,\dots,x_n)$ are homogeneous of the same degree, show that the stationary values, the turning values, and the bounding values of ϕ when $\psi = 1$ are identical with the stationary values, the turning values, and the bounding values respectively of the unrestricted function ϕ/ψ.

If ϕ, ψ are homogeneous of degrees p, q, show more generally that the stationary values, etc., of ϕ when $\psi = 1$ are identical with the corresponding values of $\phi\psi^{-p/q}$.

6. Show that the stationary values of

$$u = \sum_{r=1}^{n} a_r x_r$$

subject to the conditions $\quad \sum_{r=1}^{n} b_r x_r = 1,\qquad \sum_{r=1}^{n} x_r^2 = 1$

are the roots of the quadratic

$$\sum_r (ub_r - a_r)^2 = \sum_{r,s}(a_r b_s - a_s b_r)^2,$$

and that the greater root is a maximum, the smaller root a minimum.

Show also that the restrictive conditions do not give real values of the arguments, if $\sum b_r^2 < 1$.

7. If the variables x_r, y_s are subject to the condition

$$x_1^2+\dots+x_m^2-(y_1^2+\dots+y_n^2) = 1,$$

show that the stationary values of

$$u = (a_1 x_1^2+\dots+a_m x_m^2)-(b_1 y_1^2+\dots+b_n y_n^2)$$

are $a_1,\dots,a_m$. Show also that u may have a single minimum or a single maximum, but cannot have both. What are the conditions for the existence of a minimum?

8. If x, y are subject to the relation $lx+my+n = 0$, show that the quadratic

$$u \equiv (a,b,c,f,g,h)(x,y,1)^2$$

has the single stationary value $\quad \Delta \Big/ \dfrac{\partial \Delta}{\partial c},$

where

$$\Delta = \begin{vmatrix} a & h & g & l \\ h & b & f & m \\ g & f & c & n \\ l & m & n & 0 \end{vmatrix},$$

and that this is a maximum or minimum according as

$$\frac{\partial \Delta}{\partial c} > 0 \quad \text{or} \quad < 0.$$

Discuss the case in which $\dfrac{\partial \Delta}{\partial c} = 0$, and interpret it geometrically.

9. Show that, when x, y, z obey the condition $lx+my+nz = 1$, the quadratic

$$u \equiv (a,b,c,f,g,h)(x,y,z)^2$$

has a single stationary value, provided that

$$\begin{vmatrix} a & h & g & l \\ h & g & f & m \\ g & f & c & n \\ l & m & n & 0 \end{vmatrix} \neq 0.$$

Investigate the maxima and minima (if any) of the quadratic and discuss the case excluded above.

10. If x, y are connected by the relation

$$ax^2+2hxy+by^2 = 1,$$

show that

$$u \equiv a'x^2+2h'xy+b'y^2$$

is unbounded, unless $ab > h^2$. If, however, this condition is satisfied, show that u has a maximum and a minimum, which are also its greatest and least values, and that these extreme values are separated by both a'/a and b'/b.

11. (i) If x, y, z are connected by the relation

$$x^2+y^2+z^2 = 1,$$

show that

$$u \equiv (a,b,c,f,g,h)(x,y,z)^2$$

has three (real) stationary values, of which the greatest and the least are respectively maximum and minimum.

(ii) If x, y, z are connected by the relation

$$(a,b,c,f,g,h)(x,y,z)^2 = 1,$$

where the quadratic is invariably positive, show that

$$u = (a',b',c',f',g',h')(x,y,z)^2$$

has three (real) stationary values, of which the greatest and the least are respectively a maximum and a minimum.

(iii) Discuss the stationary, turning, and bounding values of

$$\frac{(a',b',c',f',g',h')(x,y,z)^2}{(a,b,c,f,g,h)(x,y,z)^2},$$

when the denominator is given to be an invariably positive quadratic.

12. If x, y, z are subject to the relations

$$x^2+y^2+z^2 = 1, \qquad lx+my+nz = 0,$$

show that
$$u = ax^2+by^2+cz^2$$
has two stationary values which are the roots of

$$\frac{l^2}{u-a}+\frac{m^2}{u-b}+\frac{n^2}{u-c} = 0.$$

Show that we may discriminate between maximum and minimum by considering the sign either of

$$(b-c)^2(u-a)^3/l^2+(c-a)^2(u-b)^3/m^2+(a-b)^2(u-c)^3/n^2$$

or else of
$$(a-u)(b-u)(c-u).$$
Show that the greater stationary value is a maximum, the other a minimum.

[Show that

$$\sum \frac{(b-c)^2(u-a)^3}{l^2} - \frac{(a-u)(b-u)(c-u)}{l^2m^2n^2}\left\{n^2(c-a)\frac{b-u}{c-u}-m^2(a-b)\frac{c-u}{b-u}\right\}^2$$

is divisible by $\sum l^2/(u-a)$.]

13. If x, y, z are subject to the conditions

$$x^2+y^2+z^2 = 1, \qquad lx+my+nz = p,$$

show that the stationary values of

$$u = ax^2+by^2+cz^2$$

are given by the λ-eliminant of

$$\left.\begin{aligned} \frac{l^2}{a-\lambda}+\frac{m^2}{b-\lambda}+\frac{n^2}{c-\lambda} &= \frac{p^2}{u-\lambda} \\ \frac{l^2}{(a-\lambda)^2}+\frac{m^2}{(b-\lambda)^2}+\frac{n^2}{(c-\lambda)^2} &= \frac{p^2}{(u-\lambda)^2} \end{aligned}\right\}.$$

Show also that maxima and minima may be distinguished by the sign of

$$\sum (b-c)^2(a-u)(a-\lambda)^2/l^2,$$

or alternatively by the sign of

$$(u-a)(u-b)(u-c).$$

14. If x, y, z, t are subject to the conditions

$$x^2+y^2+z^2+t^2 = 1, \qquad ax+by+cz+dt = 0, \qquad a'x+b'y+c'z+d't = 0,$$

show that the stationary values of

$$u = Ax^2+By^2+Cz^2+Dt^2$$

are the roots of the quadratic

$$\sum (cd'-c'd)^2(A-u)(B-u) = 0$$

and that these are both real. Show also that

$$(cd'-c'd)^2(B-u)+(db'-d'b)^2(C-u)+(bc'-b'c)^2(D-u)$$

and the three corresponding expressions are all positive when u is minimum and all negative when u is maximum.

15. Obtain the greatest and the least value

(i) of
$$ax^2+by^2+cz^2,$$
when the variables obey the relations
$$x^2+y^2+z^2 = 1, \qquad lx+my+nz = 0; \qquad \text{[Cf. example 12.]}$$
(ii) of
$$\frac{ax^2+by^2+cz^2}{x^2+y^2+z^2},$$
when the variables obey the relation
$$lx+my+nz = 0; \qquad \text{[Cf. example 5.]}$$
(iii) of
$$\frac{ax+by+cz}{x+y+z},$$
when the variables obey the relation
$$l^2x^2+m^2y^2+n^2z^2-2mnyz-2nlzx-2lmxy = 0,$$
and l, m, n have the same sign.

Show that the fraction in (iii) is unbounded if l, m, n have different signs.

16. If the variables $x_1,\ldots,x_n$ are subject to the condition
$$\sum_{r=1}^{m} a_r x_r = 1,$$
and n is a positive integer, show that

(i)
$$u \equiv \sum_{r=1}^{m} x_r^{2n}$$
has the minimum value
$$\left\{\sum_{r=1}^{m} a_r^{2n/(2n-1)}\right\}^{-(2n-1)}$$
and no other stationary value;

(ii)
$$v = \sum_{r=1}^{m} x_r^{2n+1}$$
has no stationary value, unless the a_r have all the same sign, but may then have 2^m distinct stationary values.

If the a_r are all positive and a_1 is the greatest of them, show that v has the single minimum
$$v = \left\{\sum_{r=1}^{m} a_r^{(2n+1)/2n}\right\}^{-2n};$$
and the single maximum
$$v = \left\{a_1^{(2n+1)/2n} - \sum_{r=2}^{m} a_r^{(2n+1)/2n}\right\}^{-2n},$$
if and only if the expression in brackets is positive.

17. If the variables x_r are restricted to be positive or zero, subject to the relation
$$\sum_{r=1}^{n} x_r^p = 1,$$
and u is defined as
$$u = \sum_{r=1}^{n} a_r x_r^q,$$
where every a_r is positive and p, $q > 1$, show that the stationary values of u are given by
$$u^{p/(p-q)} = \sum a_r^{p/(p-q)},$$

where $\sum$ is taken over *any* selection of $a_1, \ldots, a_n$. Show, however, that none of these are turning values except

$$\text{(i)}\ u = a_r \qquad (r = 1, \ldots, n),$$

$$\text{(ii)}\ u^{p/(p-q)} = \sum_{r=1}^{n} a_r^{p/(p-q)}$$

Show further that (i) are maxima and (ii) is a minimum, if $p < q$, and vice versa, if $p > q$

Discuss, in particular, the case $p = q$

18. If x, y, z are subject to the condition

$$ax + by + cz = 1,$$

show that, in general,

$$x^3 + y^3 + z^3 - 3xyz$$

has the two stationary values

$$0, \qquad (a^3 + b^3 + c^3 - 3abc)^{-1}$$

of which the first is a maximum or a minimum according as $a + b + c \gtrless 0$, but the second not a turning value

Discuss, in particular, the cases in which (i) $a + b + c = 0$, (ii) $a = b = c$

19. Obtain the stationary values of

$$x^3 + y^3 + z^3 - 3mxyz \qquad (m \neq 2),$$

when x, y, z are subject to the condition

$$x + y + z = 1,$$

and show that the symmetrical stationary value is a maximum or minimum according as $m \gtrless 2$ Show also that the other stationary values are not turning values

Show that

$$x^3 + y^3 + z^3 - 6xyz$$

has only one stationary value and no turning value

20 Obtain the stationary values of

$$x^3 + y^3 + z^3 - 3mxyz,$$

when x, y, z are subject to the condition

$$x^2 + y^2 + z^2 = 1$$

Show that the symmetrical stationary value is a maximum or a minimum according as $m \lessgtr \frac{1}{2}$, the stationary values $x = 0$, $y = z = \frac{1}{\sqrt{2}}$, etc, are maxima, if $m^2 < 1$, the stationary values $x = y = -(1 - 1/m)z$, etc, are maxima, if $m^2 > 1$

Discuss, in particular, the cases $m = \frac{1}{2}$, 1

21 Determine the stationary values and deduce the greatest and the least values of

$$\text{(i)}\ \sin\alpha + \sin\beta + \sin\gamma,$$

$$\text{(ii)}\ \sin\alpha \sin\beta \sin\gamma,$$

when $\alpha + \beta + \gamma = \omega$, a constant

Show that, in general, (i) has three distinct turning values whose sum is zero and product $-\frac{27}{4}\sin\omega$, and (ii) has three distinct turning values whose sum is $-\frac{3}{8}\sin\omega$ and product $-\frac{1}{64}\sin^3\omega$

22 If A, B, C are the angles of a plane triangle, obtain the greatest and the least values of

$$\sin mA + \sin mB + \sin mC, \qquad \cos mA + \cos mB + \cos mC,$$

$$\sin mA \sin mB \sin mC, \qquad \cos mA \cos mB \cos mC,$$

for $m = \frac{1}{2}$, 1, 2, 3 Give in each case the angles of the corresponding triangle.

In a plane triangle show that

(i) $-3\sqrt{3} \leqslant 8\sin nA \sin nB \sin nC \leqslant 3\sqrt{3}$, if n is any integer;

(ii) $\cos 2A \cos 2B \cos 2C \leqslant \cos A + \cos B + \cos C$;

(iii) $\sin 2A \sin 2B \sin 2C \leqslant \sin A \sin B \sin C \leqslant \cos \frac{1}{2}A \cos \frac{1}{2}B \cos \frac{1}{2}C$,

(iv) $\sin A \sin B \sin C \leqslant \cos \frac{1}{2}A \cos \frac{1}{2}B \cos \frac{1}{2}C$

$$\leqslant \cos \tfrac{1}{2}(B+C) \cos \tfrac{1}{2}(C+A) \cos \tfrac{1}{2}(A+B);$$

(v) $\sin 2A + \sin 2B + \sin 2C \leqslant \sin A + \sin B + \sin C \leqslant \cos \frac{1}{2}A + \cos \frac{1}{2}B + \cos \frac{1}{2}C$.

23. If A, B, C are the angles of a plane triangle, which in (iv) is acute-angled, and p, q, r are positive constants, determine the stationary values of

(i) $p\cos A + q\cos B + r\cos C$, (iii) $p\cot A + q\cot B + r\cot C$,

(ii) $\sin^p A \sin^q B \sin^r C$, (iv) $\cos^p A \cos^q B \cos^r C$,

and state the conditions to be satisfied by p, q, r in order that these values and the corresponding triangles be real.

Obtain the greatest and the least values of these four expressions.

24. Obtain the maxima and minima of

$$\text{(i)}\ \prod_{r=1}^{n} \sin \alpha_r, \qquad \text{(ii)}\ \prod_{r=1}^{n} \cos \alpha_r,$$

when $\sum_{r=1}^{n} \alpha_r = \omega$, a constant, and deduce the greatest and least values of the two expressions.

25. If l, m, n, l', m', n' are subject to the conditions

$$l^2+m^2+n^2 = 1, \qquad ll'+mm'+nn' = 0, \qquad l'^2+m'^2+n'^2 = 1,$$

and if $a > b > c$, show that

$$\tfrac{1}{2}(a-c) \geq all'+bmm'+cnn' \geq \tfrac{1}{2}(c-a)$$

26. Determine the points the sum or difference of whose tangential distances from the given circles

$$x^2+y^2+2gx+c = 0, \qquad x^2+y^2+2g'x+c = 0$$

is stationary, and show that this sum and difference have neither maxima nor minima.

27. ABC is a given triangle, P a variable point in its plane, and x, y, z are the distances AP, BP, CP. If p, q, r are given positive constants, show that $px+qy+rz$ is stationary when BC, CA, AB subtend at P angles which are the exterior angles of a triangle whose sides are proportional to p, q, r.

Obtain the least value of $px+qy+rz$ and discuss the case in which $p > q + r$.

28. Determine whether or not the origin is a maximum or minimum of the following functions:

(i) $x^2+y^3-2xy^2+x^4-y^4$,

(ii) $x^2-2xy^2+x^4-y^4$,

(iii) $x^3+y^3-2xy^2+x^4-y^4$,

(iv) $x^2-2xy+y^2-2x^3+2x^2y+2y^4$,

(v) $x^4+x^2y^3+xy^6+y^8$,

(vi) $x^2+xy^5+y^6+x^3y^4$,

(vii) $x^4+2xy^3+y^4+2x^3y^3$,

(viii) $x^2+4xy^3+2xy^4+4y^6+2xy^6+4y^7+2y^8$.

29. If the function $f(x,y,z)$, not necessarily a polynomial, is homogeneous of degree unity, show that its Hessian vanishes everywhere. Discuss its maxima and minima.

30. Show that, for the function $z = z(x,y)$ defined by the elimination of t between
$$z = xf(t)+yg(t)+h(t), \qquad xf'(t)+yg'(t)+h'(t) = 0,$$
the Hessian vanishes everywhere, and discuss the determination of the maxima and minima of z.

Illustrate by considering the following functions $z(x,y,t)$:

(i) $x\sin(t+a)+y\sin(t+b)+\sin(t-c)$,

(ii) $xt+yt^2$,

(iii) $xt+yt^2+\frac{1}{3}t^3$,

(iv) xt^2+yt^3,

(v) $xt+yt-t^2$

Show also that, for such a function $z(x,y)$, every H_n vanishes everywhere

XI

IMPLICIT FUNCTIONS. ELIMINATION

1. The nature of the problem

IN establishing the nature of functional dependence in chapter I, we distinguished, with some vehemence, between the functional relation itself and any analytical representation of that relation. The distinction is essentially one between theory and practice and, since Mathematics is an art still in its infancy, the practical aspect is nearly always the more important. To this practical aspect belongs the distinction between explicit and implicit functions, or, more exactly, between functions explicitly and implicitly defined. For, in theory, at all events, given a functional relation

$$f(x_1, x_2, .., x_n) = 0, \tag{1}$$

we are free to pick out any argument x_1 as dependent variable, i.e. to regard (1) as defining the implicit function

$$x_1 = x_1(x_2,..., x_n), \tag{2}$$

and thereupon to think of the corresponding explicit function $x_1(x_2,..., x_n)$. In practice, however, as we find from elementary and bitter experience, the passage from the implicit function to the explicit function is usually difficult and, more often than not, impossible. The problem is, of course, that of solving an equation complicated by the presence of arbitrary arguments $x_2,..., x_n$ in addition to the unknown x_1 itself, a problem, at present, generally insoluble.

It therefore becomes important to discover just how much useful information can be ascertained about the implicit function $x_1(x_2,..., x_n)$, for example, as regards existence, one-valuedness, continuity, differentiability, without needing to solve the defining equation $f(x_1, x_2,..., x_n) = 0$. In practice it will usually happen that we know not merely the *equation* $f(x_1, x_2,..., x_n) = 0$, but more generally the defining *function* $f(x_1, x_2,..., x_n)$, at any rate for values in the neighbourhood of $f = 0$, and we may therefore state our problem in the form: given information about the function of n variables $f(x_1, x_2,..., x_n)$, to obtain information about the implicit function of $n-1$ variables $x_1(x_2,..., x_n)$ defined by the equation $f(x_1, x_2,..., x_n) = 0$.

At this point we must not fail to remember that we are engaged with a theory of real functions of real variables. For it may well happen that in solving the equation (1) we encounter non-real values of x_1, although the other arguments $x_2,..., x_n$ remain real. These non-real

values must, of course, be strictly excluded from the domain of definition of the function (2). Their interpretation may belong to the theory of Functions of a Complex Variable, but it certainly has no place in the present theory.

It may, of course, even happen that the equation (1) is satisfied by no real values of $x_1,\ldots,x_n$ and therefore that the implicit function (2) does not exist in the field of real number.

If the function $f(x_1,\ldots,x_n)$ is algebraic, we know that the non-real roots x_1 occur in conjugate pairs and that, for variation of the other arguments $x_2,\ldots,x_n$, they can arise out of a pair of real roots which have passed through coincidence. Now coincidence of roots gives

$$\frac{\partial f(x_1, x_2,\ldots, x_n)}{\partial x_1} = 0. \tag{3}$$

Thus at points where $\partial f/\partial x_1 = 0$ we may expect that certain values of x_1, i.e. certain branches of the implicit function (2), will disappear (or reappear). In other words, we may expect such points to delimit the domain of some or other branch of the implicit function.

In an extreme case it may happen that the partial derivative $\partial f/\partial x_1$ vanishes 'identically', i.e. for every $x_1,\ldots,x_n$ of some region. Then in that region we know that the values of f are independent of those of x_1: in consequence the equation (1) gives no precise information about x_1, and the implicit function (2) is indeterminate. In any case the vanishing of the partial derivative $\partial f/\partial x_1$ is indicated as a danger-sign and, as we shall see, the behaviour of this partial derivative, supposed existent, is fundamental in the theory of the implicit function.

Some of the implications of this theory appear in miniature in the problem of determining the inverse of a function of one variable, i.e. in determining the function $x = x(y)$ from knowledge of the function $y = y(x)$. These are to be found in chapter III §§ 4, 5 and chapter IV § 10. In particular, it was seen that the non-vanishing of the derivative $y'(x)$ is all-important.

The following examples throw light on the character of the implicit function in certain extreme cases. Here, by a convenient change of notation, we consider the implicit function $z(x,y)$ as defined by the relation $f(x,y,z) = 0$.

(i) $f(x,y,z) \equiv (x+y+z)^{-1}$:

the implicit function z is infinite for finite values of x, y and finite only for infinite values of x, y; if we exclude 'infinite number', the implicit function does not exist;

(ii) $f(x,y,z) \equiv xyz$:

the implicit function vanishes everywhere except along the lines $x = 0 = y$, where it is indeterminate;

(iii) $f(x,y,z) \equiv x^2+y^2+z^2$:

the domain of the implicit function consists of the solitary point $x = 0 = y$;

(iv) $f(x,y,z) \equiv x^2+y^2+z^2+1$:

the implicit function does not exist in the field of the real variable;

(v) $f(x,y,z) \equiv (x^2+y^2)/(x+y+z)$:

in the field of real, finite number the implicit function exists only at the point $x = 0 = y$, where it is indeterminate.

As a less exceptional example we consider the defining function

(vi) $f(x,y,z) \equiv x^2+y^2+z^2-1$.

The implicit function may be written explicitly as $z = \sqrt{(1-x^2-y^2)}$. It exists as a real function everywhere within the circle $x^2+y^2 = 1$. As we cross the circumference of this circle, the function abruptly ceases. But we observe that

$$\frac{\partial f}{\partial z} = 2z = 2\sqrt{(1-x^2-y^2)},$$

which vanishes at all points of the circle. The domain of the implicit function is thus delimited by the points at which $\partial f/\partial z = 0$, in accordance with the principle to which we have already drawn attention.

2. Conditions sufficient for an implicit function

We come now to the enumeration of conditions sufficient to ensure the existence of an implicit function in a specified domain. For conciseness of argument and analysis we shall consider only an implicit function of two variables $z(x,y)$ defined by a relation $f(x,y,z) = 0$.

We shall suppose that, throughout some domain, $f(x,y,z)$ exists and is continuous. We shall suppose that the z-derivative $f_z(x,y,z)$ exists throughout the domain, but we shall not, at this stage, presume its continuity. We shall suppose that we know at any rate one point (x_0, y_0, z_0) at which $f(x,y,z)$ vanishes. We accordingly know at any rate one point (x_0, y_0) in the domain of the implicit function, and it is in the neighbourhood of this point that we shall seek to establish the existence of the implicit function.

We are careful to stipulate that the derivative $f_z(x,y,z)$ shall not vanish

at this point and, moreover, that we can surround z_0 by a closed interval throughout which $f_z(x_0, y_0, z)$, regarded as a function of z only, does not vanish.

With these preliminaries we are prepared for the enunciation of our main theorem:

(4) (i) *If, throughout some domain, $f(x, y, z)$ exists, is continuous, and is differentiable in z, and if (x_0, y_0, z_0) is a point of the domain such that $f(x_0, y_0, z_0)$ vanishes but $f_z(x_0, y_0, z)$ does not vanish for any z in some neighbourhood of $z = z_0$, then, throughout some neighbourhood of (x_0, y_0), the relation $f(x, y, z) = 0$ defines a bounded function $z(x, y)$.*

(ii) *If, moreover, $f_z(x, y, z)$ vanishes for no x, y, z of the domain, then the function $z(x, y)$ is one-valued.*

(iii) *If, moreover, $|f_z(x, y, z)|$ is bounded positively,† then the function $z(x, y)$ is continuous.*

To prove (i) suppose the interval (ζ, ζ') to be the neighbourhood of z_0 throughout which $f_z(x_0, y_0, z)$, regarded as a function of z, does not vanish. Then, since $f(x, y, z)$ is continuous in its three arguments, in particular $f(x_0, y_0, z)$ is a continuous function of z and so, in the closed interval (ζ, ζ'), has both a greatest value and a least value. These values must differ and must occur at the end-points of the interval, since the derivative $f_z(x_0, y_0, z)$ nowhere vanishes in the interval. It is sufficient to suppose $f(x_0, y_0, \zeta)$ the least value and $f(x_0, y_0, \zeta')$ the greatest value. Thus

$$f(x_0, y_0, \zeta) < f(x_0, y_0, z_0) < f(x_0, y_0, \zeta'),$$

i.e.

$$f(x_0, y_0, \zeta) < 0 < f(x_0, y_0, \zeta').$$

Since $f(x, y, \zeta)$ and $f(x, y, \zeta')$ are continuous functions of x, y in the domain, we can, in the domain, surround (x_0, y_0) by a closed interval‡ $(\xi, \xi'; \eta, \eta')$ throughout which $f(x, y, \zeta)$ and $f(x, y, \zeta')$ do not change sign. Hence, if (x_1, y_1) be a point of this interval,

$$f(x_1, y_1, \zeta) < 0 < f(x_1, y_1, \zeta').$$

The function $f(x_1, y_1, z)$, regarded as a function of z, has therefore each sign in the interval (ζ, ζ'), and, since it is continuous in the interval, must pass through the value zero at one or more points z_1 of the interval.

† By 'bounded positively' I mean that the function has a lower bound which is *greater than* zero. The hypothesis of (iii) is, of course, more stringent than that of (ii), since $|f_z(x, y, z)|$, not being necessarily continuous, may have zero as a lower bound without ever attaining it as a value.

‡ By the 'closed interval' $(\xi, \xi'; \eta, \eta')$ I mean, in the notation of chapter I § 4, the region

$$\xi \leqslant x \leqslant \xi', \quad \eta \leqslant y \leqslant \eta'.$$

Thus the implicit function $z(x, y)$ exists at every point (x_1, y_1) of this interval $(\xi, \xi'; \eta, \eta')$, and is bounded in the interval, since

$$\zeta < z_1 < \zeta'.$$

This establishes (i).

For (ii), since now $f_z(x, y, z)$ does not vanish in the domain, in particular $f_z(x_1, y_1, z)$, regarded as a function of z, does not vanish in (ζ, ζ'), and therefore, by Rolle's theorem, $f(x_1, y_1, z)$ has at most a single zero in the interval. In other words, z_1 is unique for given (x_1, y_1) and the implicit function $z(x, y)$ is one-valued.

For (iii) consider the values z, $z+\delta z$ of the implicit function at two near points (x, y), $(x+\delta x, y+\delta y)$ of the interval $(\xi, \xi'; \eta, \eta')$; by (ii) these values are unique. Then, by definition of the implicit function,

$$f(x, y, z) = 0 = f(x+\delta x, y+\delta y, z+\delta z),$$

so that

$$\begin{aligned} f(x, y, z)-f(x+\delta x, y+\delta y, z) \\ &= f(x+\delta x, y+\delta y, z+\delta z)-f(x+\delta x, y+\delta y, z) \\ &= \delta z\, f_z(x+\delta x, y+\delta y, z+\theta\, \delta z) \qquad (0 < \theta < 1), \end{aligned}$$

by the theorem of the mean.

Now, by the hypothesis of (iii), we have throughout the domain

$$|f_z(x, y, z)| > A > 0,$$

for some constant A. Thus

$$|f(x, y, z)-f(x+\delta x, y+\delta y, z)| > A\, |\delta z|$$

But, by the continuity of $f(x, y, z)$,

$$|\delta x|, |\delta y| < \text{some } \eta(\epsilon, x, y),$$

secures $$|f(x, y, z)-f(x+\delta x, y+\delta y, z)| < A\epsilon,$$

i.e. $$|\delta z| < \epsilon.$$

This establishes the continuity of the implicit function and proves (iii).

3. An alternative condition. Differentiability

If we now exclude functions $f(x, y, z)$ with discontinuous z-derivative, we can re-enunciate (4) in the simpler form:

(5) *If, near the point (x_0, y_0, z_0), the function $f(x, y, z)$ is continuous and possesses a continuous z-derivative $f_z(x, y, z)$, and if, at the point,*

$$f(x_0, y_0, z_0) = 0, \qquad f_z(x_0, y_0, z_0) \neq 0,$$

then in the neighbourhood of (x_0, y_0) there exists a one-valued and continuous implicit function $z(x, y)$ satisfying the relation $f(x, y, z) = 0$ and assuming the value z_0 at (x_0, y_0).

For now suppose that

$$|f_z(x_0, y_0, z_0)| = 2A > 0,$$

then, from the continuity of $f_z(x, y, z)$, we have throughout some neighbourhood of (x_0, y_0, z_0)

$$|f_z(x, y, z) - f_z(x_0, y_0, z_0)| < A$$

and so $$|f_z(x, y, z)| > A > 0.$$

The conditions of (i), (ii), (iii) of (4) are now all secured and (5) is therefore established.

As regards differentiability† of the implicit function we have the theorem:

(6) *If the conditions of* (5) *are satisfied and if, in addition,* $f(x, y, z)$ *is differentiable in its domain, then the implicit function* $z(x, y)$, *as defined in* (5), *is also differentiable.*

For, since $f(x, y, z)$ is differentiable, we may write

$$\delta f = A\,\delta x + B\,\delta y + C\,\delta z,$$

where A, B, C are functions of x, y, z, δx, δy, δz which, as $\delta x, \delta y, \delta z \to 0$, converge respectively to f_x, f_y, f_z.

If now we keep $f = 0$, we keep also $\delta f = 0$, and we have

$$-C\,\delta z = A\,\delta x + B\,\delta y, \tag{7}$$

where z now is the implicit function defined in (5). This implicit function is continuous, and therefore the convergence $\delta x, \delta y \to 0$ secures also the convergence $\delta z \to 0$.

Now, under the conditions of (5), $f_z \neq 0$ and therefore also $C \neq 0$, if δx, δy are sufficiently small. We may thus write (7) as

$$\delta z = -(A/C)\,\delta x - (B/C)\,\delta y,$$

where now $-A/C$, $-B/C$ converge respectively to $-f_x/f_z$, $-f_y/f_z$ as $\delta x, \delta y \to 0$. In other words, the implicit function $z(x, y)$ is differentiable, and its partial derivatives are given by the formulae

$$\frac{\partial z}{\partial x} = -\frac{f_x}{f_z}, \qquad \frac{\partial z}{\partial y} = -\frac{f_y}{f_z}. \tag{8}$$

These formulae we have already established in chapter VI § 11, but then only by presupposing the existence of the implicit function.

† Of course, in the strict sense of chapter VI § 2 (8), and so throughout this section.

We proceed in this way and prove in extension of (6) that

(9) *With the conditions of* (5), *if* $f(x,y,z)$ *is twice differentiable, then the implicit function* $z(x,y)$ *is also twice differentiable.*

For, since $f(x,y,z)$ is now twice differentiable, the partial derivatives f_x, f_y, f_z are also differentiable in x, y, z.

Now, by chapter VI (14), differentiable functions of differentiable functions are themselves differentiable, and therefore, if the argument z is replaced by the implicit function $z(x,y)$, the partial derivatives f_x, f_y, f_z become differentiable functions of x, y. Since zeros of f_z have throughout been excluded, it follows from (8) that $\partial z/\partial x$, $\partial z/\partial y$ are differentiable in x, y; in other words, that the implicit function $z(x,y)$ is twice differentiable.

By a similar argument we can show that $z(x,y)$ is differentiable n times, if $f(x,y,z)$ is differentiable n times.

4. Condition for an analytic implicit function

For an analytic function of several variables we shall adopt the definition outlined in the concluding section of chapter VIII: we shall say that $f(x,y,z)$ is analytic in an interval in which every partial derivative exists and satisfies an inequality

$$\left|\frac{\partial^{m+n+p}f}{\partial x^m \partial y^n \partial z^p}\right| < m!\,n!\,p!\,AB^{m+n+p},$$

where A, B are positive numbers independent of x, y, z, m, n, p.

In the sense of this definition we proceed to prove the theorem:

(10) *With the conditions of* (5) *the implicit function* $z(x,y)$ *is analytic, if the defining function* $f(x,y,z)$ *is itself analytic.*

By the results of the last section, every derivative of $z(x,y)$ exists, since every derivative of $f(x,y,z)$ exists. Now write

$$\frac{\partial^{m+n}z}{\partial x^m \partial y^n} = \sum_{s=1}^{2m+2n-1}\left(-\frac{\partial f}{\partial z}\right)^{-s} Z_{mns}, \tag{11}$$

$$f_{pqr} = p!\,q!\,r!\,u_{pqr}. \tag{12}$$

By the formulae of chapter VI § 11, we can then write Z_{mns} as a polynomial function of arguments u_{pqr}. Comparison of dimensions in f, x, y, z respectively shows that this polynomial $Z_{mns}(u_{pqr})$ is homogeneous of degree s in these arguments u and isobaric of weights m, n, $s-1$ in the respective suffixes p, q, r.

Differentiation of (11) with respect to x gives

$$\sum_{s=1}^{2m+2n+1}\left(-\frac{\partial f}{\partial z}\right)^{-s} Z_{m+1,n,s}$$
$$= \sum_{s=1}^{2m+2n-1}\left\{\left(-\frac{\partial f}{\partial z}\right)^{-s}\left(\frac{\partial Z_{mns}}{\partial x}+\frac{\partial z}{\partial x}\frac{\partial Z_{mns}}{\partial z}\right)+\right.$$
$$\left.+s\left(-\frac{\partial f}{\partial z}\right)^{-s-1}\left(\frac{\partial^2 f}{\partial x\partial z}+\frac{\partial z}{\partial x}\frac{\partial^2 f}{\partial z^2}\right)Z_{mns}\right\}$$
$$= \sum_{s=1}^{2m+2n-1}\left\{\left(-\frac{\partial f}{\partial z}\right)^{-s}\frac{\partial Z_{mns}}{\partial x}+\left(-\frac{\partial f}{\partial z}\right)^{-s-1}\left(\frac{\partial f}{\partial x}\frac{\partial Z_{mns}}{\partial z}+s\frac{\partial^2 f}{\partial x\partial z}Z_{mns}\right)+\right.$$
$$\left.+s\left(-\frac{\partial f}{\partial z}\right)^{-s-2}\frac{\partial f}{\partial x}\frac{\partial^2 f}{\partial z^2}Z_{mns}\right\},$$

by use of the formulae of chapter VI § 11.

We therefore have for $Z_{m+1,n,s}$ the recurrence-formula

$$Z_{m+1,n,s} = \frac{\partial Z_{mns}}{\partial x}+\frac{\partial f}{\partial x}\frac{\partial Z_{m,n,s-1}}{\partial z}+(s-1)\frac{\partial^2 f}{\partial x\partial z}Z_{m,n,s-1}+$$
$$+(s-2)\frac{\partial f}{\partial x}\frac{\partial^2 f}{\partial z^2}Z_{m,n,s-2}. \qquad (13)$$

Since only positive signs occur on the right of this recurrence-formula, it is clear, by induction, that only positive signs will occur in the polynomials Z_{mns}. Let c_{mns} denote the sum of the coefficients in the polynomial Z_{mns}: then c_{mns} is obtained from Z_{mns} by writing unity for every u_{pqr}.

Now, by (12),

$$\frac{\partial Z_{mns}}{\partial x} = \sum (p+1)u_{p+1,q,r}\frac{\partial Z_{mns}}{\partial u_{pqr}}, \qquad (14)$$

where the summation is taken over all the arguments u_{pqr} of Z_{mns}. Again, by the formulae (47), (54) of chapter VI for homogeneous and for isobaric functions, we have, in virtue of the dimensional properties of Z_{mns} noted above,

$$\sum p u_{pqr}\frac{\partial Z_{mns}}{\partial u_{pqr}} = mZ_{mns},$$
$$\sum u_{pqr}\frac{\partial Z_{mns}}{\partial u_{pqr}} = sZ_{mns},$$

i.e.
$$\sum (p+1)u_{pqr}\frac{\partial Z_{mns}}{\partial u_{pqr}} = (m+s)Z_{mns}.$$

Hence, by putting unity for every u_{pqr} in this equation and in (14), we

see that the sum of the coefficients in $\partial Z_{mns}/\partial x$ equals the sum of the coefficients in $(m+s)Z_{mns}$, i.e. equals $(m+s)c_{mns}$.

By a similar argument we see that the sum of the coefficients in $\partial Z_{mns}/\partial z$ is equal to $(2s-1)c_{mns}$. On picking out the sums of the coefficients in the recurrence-formula (13), we have the recurrence-formula for c_{mns}, namely

$$c_{m+1,n,s} = (m+s)c_{mns}+(3s-4)c_{m,n,s-1}+2(s-2)c_{m,n,s-2}. \tag{15}$$

Now, by (11),

$$\left(\frac{\partial f}{\partial z}\right)^{2m+2n-1}\frac{\partial^{m+n}z}{\partial x^m \partial y^n} = -\sum_{s=1}^{2m+2n-1}\left(-\frac{\partial f}{\partial z}\right)^{2m+2n-s-1} Z_{mns},$$

and we can therefore write the expression on the left as a polynomial in u_{pqr}. For C_{mn}, the sum of the *absolute* values of its coefficients, we have

$$C_{mn} = \sum_{s=1}^{2m+2n-1} c_{mns}.$$

From the recurrence-formula (15) we have

$$\sum_{s=1}^{2m+2n+1} c_{m+1,n,s} = \sum_{s=1}^{2m+2n-1} (6s+m-1)c_{mns}$$
$$< (13m+12n-7)\sum_{s=1}^{2m+2n-1} c_{mns},$$

since every c_{mns} is positive. Thus

$$C_{m+1,n} < (13m+12n-7)C_{mn}$$
$$< 13(m+n+1)C_{mn};$$

similarly $\qquad C_{m,n+1} < 13(m+n+1)C_{mn}.$

Hence, by induction, $\quad C_{mn} < 13^{m+n}(m+n)!\,C_{00},$

i.e. $\qquad < 13^{m+n}(m+n)!$

Now, if $f(x,y,z)$ is analytic,

$$|u_{pqr}| < AB^{p+q+r},$$

where A, B are some constants independent of x, y, z, p, q, r. But

$$\frac{\partial^{m+n}z}{\partial x^m \partial y^n}\left(\frac{\partial f}{\partial z}\right)^{2m+2n-1}$$

is isobaric of weights m, n, $2(m+n-1)$ in the suffixes p, q, r of u_{pqr} and homogeneous of degree $2m+2n-1$ in u_{pqr} itself, as we see by comparing dimensions in x, y, z, f respectively; and the sum of the absolute values of its coefficients is C_{mn}. Hence

$$\left|\frac{\partial^{m+n}z}{\partial x^m \partial y^n}\right| \left|\frac{\partial f}{\partial z}\right|^{2m+2n-1} < C_{mn}A^{2m+2n-1}B^{3m+3n-2}.$$

Again within the domain of the implicit function $z(x,y)$, as defined by (5),

$$\left|\frac{\partial f}{\partial z}\right| > \text{some positive } H^{-1}$$

and so
$$\left|\frac{\partial^{m+n}z}{\partial x^m \partial y^n}\right| < (m+n)!(AB^2H^{-1})(13A^2B^3H^2)^{m+n},$$

which proves $z(x,y)$ an analytic function.

5. Analytic prolongation

To sum up the theorems of the preceding sections we may say, in loose language, that the implicit function $z(x,y)$ exists and behaves (as regards differentiation and so forth) like the defining function $f(x,y,z)$ except near zeros of $f_z(x,y,z)$. This, of course, is not at variance with what both general considerations and particular examples had led us to expect.

These zeros of $f_z(x,y,z)$ are, of course, points in the (x,y,z)-plane, whereas the implicit function is being examined in the (x,y)-plane. The points of possible singularity are therefore strictly the points (if any) in the (x,y)-plane at which simultaneously

$$f(x,y,z) = 0 = f_z(x,y,z).$$

The aggregate of such points we call the 'z-discriminant' of $f(x,y,z)$, by a natural extension of the elementary use of the term. Thus we may say that the implicit function exists and is well-behaved except possibly in the neighbourhood of points of the discriminant.

Actually, in the analysis of the foregoing sections we have proved the existence of the implicit function $z(x,y)$ and discussed its properties over an interval $(\xi,\xi';\eta,\eta')$. This interval contains the initial point (x_0,y_0), at which by hypothesis the implicit function is presumed to exist, and it just falls short of the nearest point of the discriminant. Now we may be sure that, in general, this rectangle is not the complete domain of existence of the implicit function. A sufficient example to the contrary is afforded by the equation

$$f(x,y,z) = x^2+y^2+z^2-1 = 0, \tag{16}$$

which defines the implicit function

$$z(x,y) = \sqrt{(1-x^2-y^2)},$$

and clearly defines it through the interior of the circle $x^2+y^2 = 1$.

Actually the interval $(\xi,\xi';\eta,\eta')$ first enters the analysis in the proof

of (4) as a region throughout which $f(x, y, \zeta)$, $f(x, y, \zeta')$ retain the respective signs of $f(x_0, y_0, \zeta)$, $f(x_0, y_0, \zeta')$ and the choice of a rectangle as such a unit region arises merely from custom and convenience.

It is, however, immaterial precisely what sort of region we adopt as our conventional unit, since, in general, any point of the domain can ultimately be reached by a process of 'analytic prolongation' similar to that already described for the analytic function.†

The process is simple. With (x_0, y_0, z_0) as point of departure we have been able to determine a domain $(\xi, \xi'; \eta, \eta')$ for the implicit function. Taking any point (x_1, y_1) of this interval together with z_1 the value there of the implicit function, we have in the (x, y, z)-plane, a point (x_1, y_1, z_1), at which the conditions of (5) are satisfied and which thus gives us a domain $(\xi_1, \xi_1'; \eta_1, \eta_1')$ for the implicit function. This new interval will overlap the old in some neighbourhood of (x_1, y_1), but, in general, will not lie entirely within it. We have then been able to extend or 'prolong' the domain of the implicit function. Again, from this extended domain we can select a new point (x_2, y_2) which, in turn, gives us a new interval $(\xi_2, \xi_2'; \eta_2, \eta_2')$, in general prolonging the domain and so on.

As a rule we can get in this way from any initial point (x_0, y_0) to any other point (x, y) of the domain of the implicit function by a finite sequence of intervals $(\xi_r, \xi_r'; \eta_r, \eta_r')$. The process is that of bill-posting over an irregular hoarding. We can, if we wish, visualize the exact domain of the implicit function as being the limit of an infinite sequence of such intervals, but this problem of the precise determination of this domain is a more practical problem with whose technique we are not now concerned.

It may well happen, of course, that the discriminant splits up the (x, y)-plane, and with it the domain of the implicit function, into disconnected areas. Consider, for instance, the equation

$$f(x, y, z) \equiv (x^2+y^2)^2-4(x^2+y^2+z^2)+3 = 0 \tag{17}$$

which defines the implicit function

$$z = \tfrac{1}{2}\sqrt{\{(x^2+y^2-1)(x^2+y^2-3)\}}.$$

The domain of this function consists of the interior of the circle $x^2+y^2 = 1$ and the exterior of the circle $x^2+y^2 = 3$, two disconnected areas, the discriminant being composed of the circumferences of these two circles. Clearly we cannot now pass from any one point of the domain to any other: we must know a distinct point of departure for each distinct region.

† Chapter VIII § 6.

6. Branches and branch-points

There is one other point to be considered. The theory of the implicit function which we have constructed presents that function as a one-valued, continuous function, but we know from experience that, as often as not, the implicit function is many-valued, as witness the last two examples (16), (17). We have therefore to find room in our theory for this many-valuedness. The solution is, of course, obvious. If the implicit function is, in actual fact, many-valued, let us say n-valued, then, associated with the chosen initial point of its domain (x_0, y_0), we have the n values $z_{01}, \ldots, z_{0n}$ of z_0. There are thus n distinct points of departure

$$(x_0, y_0, z_{01}), \ldots, (x_0, y_0, z_{0n})$$

in the (x, y, z)-plane, which, by our theory, give rise to n apparently distinct one-valued, continuous functions. These are the n 'branches' of the implicit function to which attention has often been directed. Evidently then, our present theory exhibits these branches to all intents and purposes as starkly independent, one-valued functions. This is as it should be, for the many-valued function, so split up, becomes at once admissible to the whole theory of the one-valued function, and the distinction between the many-valued function and the one-valued function ceases to be formidable.

We may ask how far these branches retain their individual identity. If the defining equation is polynomial, we know that two roots of $f(x, y, z) = 0$, regarded as an equation in z, coincide only if the equation $f_z(x, y, z) = 0$ is also satisfied, i.e. that two branches of the implicit function coincide only at a point of the discriminant. More generally for any function, if $z(x, y)$, $z'(x, y)$ are two branches of the implicit function, i.e. if simultaneously

$$f(x, y, z) = 0 = f(x, y, z'),$$

then the theorem of the mean gives

$$f_z(x, y, Z) = 0,$$

where Z lies between z, z'. Hence, if z, z' move up to coincidence as (x, y) moves up to some point in the (x, y)-plane, then at such a point $f_z(x, y, z) = 0$, since f_z is continuous. In other words, the point is a point of the discriminant.

Points at which branches become coincident are called *branch-points* and we have therefore found that

(18) *Branch-points are points of the discriminant.*

As an example consider the function (16)

$$f(x,y,z) \equiv x^2+y^2+z^2-1.$$

The implicit function has the two branches

$$z_1 = +(1-x^2-y^2)^{\frac{1}{2}}, \qquad z_2 = -(1-x^2-y^2)^{\frac{1}{2}},$$

distinguished only by their signs. Since z_1 is always positive and z_2 is always negative, these branches can become confused only by passage through zero, i.e. through points in the (x,y)-plane at which $x^2+y^2 = 1$. These are exactly the points of the discriminant, which is here a ring of branch-points bounding the domain of the implicit function externally. Within this domain the two branches retain their separate identity.

Or again, consider the equation

$$f(x,y,z) \equiv x^2+y^2-z^2 = 0 \tag{19}$$

which defines the two-valued implicit function

$$z = \pm(x^2+y^2)^{\frac{1}{2}}.$$

The discriminant reduces to the solitary point $x = 0 = y$, which is the only branch-point, and the implicit function is defined throughout the (x,y)-plane. Its two branches remain distinct except for passage through the solitary branch-point.

In (19), of course, z is merely the polar coordinate r, thinly disguised. For the polar coordinate θ consider the equation

$$f(x,y,z) \equiv x \sin z - y \cos z = 0 \tag{20}$$

which defines the many-valued implicit function

$$z = \tan^{-1}(y/x).$$

The z-discriminant is given by

$$x \cos z + y \sin z = 0 = x \sin z - y \cos z,$$

whence $$x^2+y^2 = 0.$$

The discriminant again reduces to the single point $x = 0 = y$, which is the only branch-point. But, if x, y be expressed in polar coordinates in the (x,y)-plane, we see at once that we may write $z = \theta+n\pi$.

Suppose that at the initial point (x_0, y_0, z_0) we have $\theta = \theta_0$; then the values of z_0 which define the infinitely many branches are

$$\theta_0,\ \theta_0 \pm \pi,\ \theta_0 \pm 2\pi, \ldots \text{ to } \infty,$$

say $$z_{0n} \equiv \theta_0 + n\pi,$$

where n is an integer of either sign, or zero. If now, starting from (x_0, y_0), we describe a circle counter-clockwise about the branch-point, the polar coordinate θ has been increased by 2π on our return to (x_0, y_0). In other words, any branch $z_{0n}(x, y)$, by travelling round this circle, has become transformed into the neighbouring branch $z_{0,n+2}(x, y)$. In such a case the individual branches no longer preserve their individuality, even though the branch-point be screened from the domain of the implicit function. To define the implicit function without ambiguity it is no longer sufficient merely to specify the initial (x_0, y_0, z_0) without ambiguity; we must also specify the route that has brought us from (x_0, y_0) to the point (x, y), or alternatively we must interpose some obstacle into the domain of the implicit function which will prevent circumnavigation of the branch-point. A happier alternative is Riemann's, in which we replace the simple (x, y)-plane by an (x, y)-plane of many sheets, matching the many-valued z with a many-valued (x, y). The discussion of these alternatives would lead us beyond our proper limits. It is sufficient here to point out that the fundamental theorem (5) still stands. The shortcoming which we *have referred to here belongs only to the process of analytic prolongation*, and not to the implicit function defined in the unit interval $(\xi, \xi'; \eta, \eta')$.

We may add that, if there is no real discriminant, the implicit function is necessarily one-valued. For, given (x, y), if $f_z(x, y, z)$ vanish for no z, then, by Rolle's theorem, $f(x, y, z)$, regarded as a function of z, vanishes at most once, and the implicit function $z(x, y)$ is therefore one-valued.

7. Systems of implicit functions

We can extend the foregoing theory to the case of n equations $f_1 = 0, \ldots, f_n = 0$ defining n implicit functions $z_1, \ldots, z_n$. The decisive part which has hitherto been played in the theory by the z-derivative $\partial f/\partial z$ is now taken by the z-Jacobian of the n defining functions, namely

$$\frac{\partial(f_1, \ldots, f_n)}{\partial(z_1, \ldots, z_n)}.$$

It will be sufficient still to suppose that the field of the implicit functions themselves is a field of two variables (x, y). In these terms we can state the theorem corresponding to (5):

(21) *If in the neighbourhood of some point* $(x_0, y_0, z_{10},..., z_{n0})$ *the n functions and their* n^2 *z-derivatives*

$$f_r(x, y, z_1,..., z_n) \quad (r = 1,..., n), \qquad \partial f_r/\partial z_s \quad (r, s = 1,..., n)$$

all exist and are continuous, and if at the point itself we have simultaneously

$$f_r = 0 \quad (r = 1,..., n), \qquad \partial(f_1,..., f_n)/\partial(z_1,..., z_n) \neq 0,$$

then near the point (x_0, y_0) *there exist the n one-valued continuous functions*

$$z_r(x, y) \qquad (r = 1,..., n)$$

which simultaneously satisfy the n equations

$$f_r(x, y, z_1,..., z_n) = 0 \qquad (r = 1,..., n).$$

The proof is inductive. The theorem evidently reduces to (5), if $n = 1$, and we therefore begin by supposing that the theorem holds for $n-1$ such equations defining $n-1$ implicit functions.

Now since the z-Jacobian of the n functions $f_1,..., f_n$ does not vanish at $(x_0, y_0, z_{10},..., z_{n0})$, it is clear that the elements of its first row, namely

$$\partial f_1/\partial z_r \qquad (r = 1,..., n),$$

cannot all vanish at this point. Let us suppose that $\partial f_1/\partial z_1$ does not vanish there. Then, by (5), we can solve the equation $f_1 = 0$ for z_1 and determine in the neighbourhood of $(x_0, y_0, z_{20},..., z_{n0})$ the one-valued continuous function

$$z_1 = z_1(x, y, z_2,..., z_n). \tag{22}$$

Moreover, by (6), this function is differentiable in $z_2,..., z_n$, since f_1 is differentiable in $z_1, z_2,..., z_n$ and $\partial f_1/\partial z_1 \neq 0$ near the initial point; in addition, these derivatives

$$\partial z_1/\partial z_r \qquad (r = 2,..., n)$$

are continuous, since the partial derivatives $\partial f_1/\partial z_r$ are themselves continuous.

Substitution for z_1 from (22) in the $n-1$ functions $f_2,..., f_n$ gives the $n-1$ functions of x, y, $z_2,..., z_n$ only

$$F_r(x, y, z_2,..., z_n) \equiv f_r(x, y, z_1, z_2,..., z_n) \qquad (r = 2,..., n). \tag{23}$$

We proceed to show that these $n-1$ functions F_r satisfy the conditions of (21), namely that

(i) they are continuous in some neighbourhood of $(x_0, y_0, z_{20},..., z_{n0})$;

(ii) their z-derivatives $\partial F_r/\partial z_s$ $(r, s = 2,..., n)$ are continuous in this neighbourhood;

(iii) they vanish simultaneously at $(x_0, y_0, z_{20},..., z_{n0})$;

(iv) their Jacobian $\partial(F_2,..., F_n)/\partial(z_2,..., z_n)$ does not vanish at this point.

Of these, (i) is immediate, since, by definition, $F_2,\ldots, F_r$ are (one-valued) continuous functions of $x, y, z_2,\ldots, z_n$ and of z_1, which is itself a one-valued, continuous function of $x, y, z_2,\ldots, z_n$.

Again, by the formula of the total differential, chapter VI (18),

$$\frac{\partial F_r}{\partial z_s} = \frac{\partial f_r}{\partial z_s} + \frac{\partial f_r}{\partial z_1}\frac{\partial z_1}{\partial z_s} \qquad (r, s = 2,\ldots, n). \tag{24}$$

The formula is valid, since, in the neighbourhood considered, the partial derivatives $\partial f_r/\partial z_s$ are, by hypothesis, continuous and we have shown that the partial derivatives $\partial z_1/\partial z_s$ exist.

Moreover, since the partial derivatives $\partial f_r/\partial z_s$, $\partial z_1/\partial z_s$ are all continuous in the neighbourhood in question, the partial derivatives $\partial F_r/\partial z_s$ are themselves continuous, which proves (ii).

Taking (iii) as obvious from (23), we proceed to prove (iv). If in the Jacobian $\partial(f_1,\ldots,f_n)/\partial(z_1,\ldots, z_n)$ we multiply the first column by $\partial z_1/\partial z_s$ and add it to the sth column ($s = 2,\ldots, n$), the element in the rth row and sth column ($r, s = 2,\ldots, n$) is then $\partial F_r/\partial z_s$ by (24). The constituent of the first row and sth column ($s = 2,\ldots, n$) is

$$\frac{\partial f_1}{\partial z_s} + \frac{\partial f_1}{\partial z_1}\frac{\partial z_1}{\partial z_s},$$

which is zero by definition of z_1. The determinant can then be reduced by the first row and we get

$$\frac{\partial(f_1,\ldots,f_n)}{\partial(z_1,\ldots,z_n)} = \frac{\partial f_1}{\partial z_1}\frac{\partial(F_2,\ldots,F_n)}{\partial(z_2,\ldots,z_n)}.$$

This gives (iv), since, by hypothesis, the Jacobian on the left is not zero at $(x_0, y_0, z_{10},\ldots, z_{n0})$.

Thus the $n-1$ functions $F_2,\ldots, F_n$ satisfy the conditions of the theorem, and therefore, by the inductive hypothesis, the $n-1$ equations

$$F_r = 0 \qquad (r = 2,\ldots, n)$$

define in some neighbourhood of (x_0, y_0) the $n-1$ one-valued, continuous implicit functions

$$z_s(x, y) \qquad (s = 2,\ldots, n).$$

Substitution in (22) adds to these the nth implicit function $z_1(x, y)$, which is continuous in the same neighbourhood, since continuous functions of continuous functions are themselves continuous. Thus, on the hypothesis that the theorem holds for $n-1$ equations, it has been shown to hold also for n, and therefore, since it is true for $n = 1$, it is established inductively.

8. Necessity of the conditions?

We may ask how far the conditions, shown in (5) and (21) to be sufficient for the existence of the implicit function, are also necessary. Let us take continuity and differentiability for granted and confine attention to the characteristic conditions

$$\frac{\partial f}{\partial z} \neq 0, \qquad \frac{\partial(f_1,\dots,f_n)}{\partial(z_1,\dots,z_n)} \neq 0.$$

The second of these, as we have seen, can be made to depend on the first, and we shall therefore consider only a single defining function $f(x,y,z)$.

Comparison of the pair of equations

$$x^2+y^2+z^2 = 0, \qquad x^2+y^2-z^2 = 0 \tag{25}$$

at the initial point $(0,0,0)$ reminds us that a necessary condition for the existence of the implicit function is that the initial point be neither a maximum nor a minimum of the defining function. For, if it be a turning-point, then, in the neighbourhood of (x_0,y_0,z_0),

either $$f(x,y,z) > f(x_0,y_0,z_0) = 0,$$

or $$f(x,y,z) < f(x_0,y_0,z_0) = 0,$$

unless, exceptionally the initial point lie on a line of maxima and minima. From this point of view the condition $f_z \neq 0$ is effective in securing that the initial point shall not be a stationary point and therefore certainly not a turning point.

If now $f_z(x_0,y_0,z_0) = 0$, but the implicit function $z(x,y)$ still exists, we have no longer the assurance that it will be continuous at (x_0,y_0). Thus

$$z(x^2+y^2) = 2xy \tag{26}$$

defines the implicit function

$$z = \frac{2xy}{x^2+y^2},$$

which is discontinuous at $(0,0)$. Nor need the implicit function be one-valued near the initial point, as is shown by the second example of (25). On the other hand, the implicit function, in the real field, is not necessarily many-valued, nor discontinuous as is seen from the defining equation

$$z^3 = x^3+y^3. \tag{27}$$

More seriously, the implicit function may have a faulty domain of definition. Thus

$$z^3 \sin\frac{1}{x} = y \tag{28}$$

defines an implicit function $z(x,y)$ except along the lines $x = (n\pi)^{-1}$, where z becomes infinite. Again

$$z^2(x^2+y^2)\sin\frac{1}{x^2+y^2} = x^4+y^4 \tag{29}$$

defines an implicit function, only in the infinite set of circular rings

$$\frac{1}{\pi} \leqslant x^2+y^2, \qquad \frac{1}{(2n+1)\pi} \leqslant x^2+y^2 \leqslant \frac{1}{2n\pi} \quad (n = 1,2,\ldots).$$

The above examples are enough to show the virtues of the condition $f_z \neq 0$, even though it is only a sufficient but not an exact condition.

9. Elimination

We pass naturally from the theory of implicit functions to that of elimination. Let us begin with the example of a set of linear equations

$$\left.\begin{aligned} a_1x+b_1y+c_1z+d_1 &= 0 \\ a_2x+b_2y+c_2z+d_2 &= 0 \\ a_3x+b_3y+c_3z+d_3 &= 0 \end{aligned}\right\}. \tag{30}$$

If (i) the Jacobian in x, y, z of the set, i.e. the determinant

$$\begin{matrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{matrix}$$

does not vanish, we can solve the equations in x, y, z, the solution being unique. But, if (ii) this determinant does vanish, then the three equations fail to determine x, y, z: they are either inconsistent or redundant. For consistency we need that the set of determinants

$$\begin{matrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{matrix} \tag{31}$$

vanish.† If this is the case, the three equations (30) are then redundant, being equivalent at most to two independent equations, so that x, y, z cannot now be determined from them with less than one degree of freedom.

If we shift the emphasis from x, y, z to the other elements of the equations, we may say that in case (i), when the Jacobian does not vanish, the system of equations (30) imposes no restriction on any element except x, y, z themselves. But in case (ii), when the Jacobian vanishes, we can deduce from (30) information, namely (31), about the

† The condition is equivalent to two independent conditions, of which the vanishing of the Jacobian counts one.

elements other than x, y, z or, as we may say, x, y, z have been *eliminated* from the three equations and (31) is their eliminant.†

We see this more starkly, if we begin with the strictly homogeneous equations

$$\left.\begin{aligned} a_1x+b_1y+c_1z &= 0 \\ a_2x+b_2y+c_2z &= 0 \\ a_3x+b_3y+c_3z &= 0 \end{aligned}\right\}. \tag{32}$$

For now, in general, we can solve these in x, y, z getting $x = 0$, $y = 0$, $z = 0$. But, if the Jacobian vanish, the three equations are redundant, and x, y, z are determined with at least one degree of freedom, i.e. at most only their ratios are known. In this case the eliminant of the three equations is the Jacobian itself.

It may make the distinction a little clearer, if we say that, humanly speaking, when we are interested in x, y, z, then we prefer that the Jacobian *should not* vanish: when we are not interested in x, y, z, we prefer that the Jacobian *should* vanish.

10. Conditions for the existence of an eliminant

It follows, then, from theorem (21) that the n equations

$$f_r(z_1,\ldots,z_n,\ldots) = 0 \qquad (r = 1,\ldots,n)$$

will not have an eliminant in $z_1,\ldots,z_n$ unless, granted differentiability and the like, the Jacobian

$$\frac{\partial(f_1,\ldots,f_n)}{\partial(z_1,\ldots,z_n)}$$

vanishes whenever the given equations are satisfied. For, in the neighbourhood of a point where the Jacobian does not vanish, we can determine the implicit functions $z_1,\ldots,z_n$ without imposing any condition on the remaining elements of the n functions $f_1,\ldots,f_n$, i.e. no eliminant exists in such a neighbourhood. We may therefore state the theorem:

(33) *If $z_1,\ldots,z_n$ can be eliminated from the n equations $f_1 = 0,\ldots, f_n = 0$, then it is necessary that*

$$\frac{\partial(f_1,\ldots,f_n)}{\partial(z_1,\ldots,z_n)} = 0$$

whenever the n equations are satisfied.

This condition is not, however, sufficient to secure the existence of an eliminant. For consider the pair of equations

$$f_1 \equiv 2z_1^2+z_2-a = 0 \qquad f_2 \equiv z_1^2+z_2-a = 0. \tag{34}$$

† In what follows I shall not distinguish between the eliminant as an *expression* e.g. (31), and the eliminant as a *condition*, i.e. the expression equated to zero. The distinction is important in certain connexions, for instance in the study of algebraic eliminants.

The Jacobian is
$$\frac{\partial(f_1,f_2)}{\partial(z_1,z_2)} = 2z_1.$$

This vanishes, if $f_1-f_2 = 0$, i.e. certainly, if $f_1 = 0$, $f_2 = 0$. But the pair of equations (34) have the solution
$$z_1 = 0, \qquad z_2 = a,$$
and there is no eliminant.

In contrast to (33), which is necessary but not sufficient for the existence of an eliminant, we can give the following condition which is sufficient but not necessary:

(35) *A sufficient condition that the n equations*
$$f_1 = 0, \quad ..., \quad f_n = 0$$
have an eliminant in $z_1,...,z_n$ is that the Jacobian
$$J = \frac{\partial(f_1, ...f_n)}{\partial(z_1,..., z_n)}$$
vanish whenever the $n-1$ equations
$$f_2 = 0, \quad ..., \quad f_n = 0$$
are satisfied, provided also that those first minors of J which do not involve f_1, i.e. the set† of Jacobians
$$\frac{\partial(f_2,...,f_n)}{\partial((z_1,..., z_n))},$$
do not all vanish when $f_2 = 0$, ..., $f_n = 0$.

To prove this let us suppose that
$$\frac{\partial(f_2,...,f_n)}{\partial(z_2,..., z_n)}$$
is a minor of J that does not vanish when $f_2,...,f_n$ vanish. Then, by (21), $z_2,..., z_n$ can be developed in some suitable domain as functions of z_1 and the other elements of $f_1,...,f_n$, say
$$z_r = \phi_r(z_1,...) \qquad (r = 2,..., n).$$

Substitution in $f_1 = 0$ gives
$$F(z_1) = f_1\{z_1, \phi_2(z_1,...),..., \phi_n(z_1,...)\} = 0.$$

† I use this notation to represent the set of n determinants
$$\left\| \begin{matrix} \frac{\partial f_2}{\partial z_1} & . & . & . & \frac{\partial f_2}{\partial z_n} \\ . & . & . & . & . \\ \frac{\partial f_n}{\partial z_1} & . & . & . & \frac{\partial f_n}{\partial z_n} \end{matrix} \right\|$$

There is certainly an eliminant, then, if z_1 drops out of this relation, i.e. if

$$\frac{\partial F}{\partial z_1} = 0.$$

Now
$$\frac{\partial F}{\partial z_1} = \frac{\partial f_1}{\partial z_1} + \frac{\partial f_1}{\partial \phi_2}\frac{\partial \phi_2}{\partial z_1} + \dots + \frac{\partial f_1}{\partial \phi_n}\frac{\partial \phi_n}{\partial z_1}. \tag{36}$$

But, by definition of $\phi_2,\dots,\phi_n$, we have identically

$$f_r\{z_1, \phi_2(z_1,\dots),\dots,\phi_n(z_1,\dots)\} = 0 \qquad (r = 2,\dots,n),$$

and so
$$\frac{\partial f_r}{\partial z_1} + \frac{\partial f_r}{\partial \phi_2}\frac{\partial \phi_2}{\partial z_1} + \dots + \frac{\partial f_r}{\partial \phi_n}\frac{\partial \phi_n}{\partial z_1} = 0 \qquad (r = 2,\dots,n).$$

If we solve these equations in

$$\frac{\partial \phi_2}{\partial z_1}, \quad \dots, \quad \frac{\partial \phi_n}{\partial z_1}$$

and substitute in (36), we get

$$\frac{\partial(f_2,\dots,f_n)}{\partial(\phi_2,\dots,\phi_n)}\frac{\partial F}{\partial z_1} = \frac{\partial(f_1, f_2,\dots,f_n)}{\partial(z_1, \phi_2,\dots,\phi_n)}. \tag{37}$$

Now $z_2 = \phi_2$, ..., $z_n = \phi_n$ are, by definition, consequences of $f_2 = 0$, ..., $f_n = 0$. Accordingly, when $f_2 = 0$, ..., $f_n = 0$, we may write (37) as

$$\frac{\partial(f_2,\dots,f_n)}{\partial(z_2,\dots,z_n)}\frac{\partial F}{\partial z_1} = \frac{\partial(f_1, f_2,\dots,f_n)}{\partial(z_1, z_2,\dots,z_n)}. \tag{38}$$

But, by hypothesis, the first Jacobian in (38) does not vanish, and the second Jacobian does vanish, when $f_2 = 0,\dots, f_n = 0$. Thus $\partial F/\partial z_1$ vanishes, and consequently z_1 does not appear in the relation $F = 0$, which is therefore an eliminant of $f_1,\dots,f_n$ with respect to $z_1,\dots,z_n$, and the theorem is proved.

The conditions of (35), though sufficient for an eliminant, are not necessary. For consider the pair of equations

$$\left.\begin{aligned} f_1 &\equiv z_1^2 - z_2^2 - 2az_1 + 2bz_2 - 1 &&= 0 \\ f_2 &\equiv z_1^2 - z_2^2 - (a+b)(z_1 - z_2) - 1 - a^2 + b^2 &&= 0 \end{aligned}\right\}. \tag{39}$$

We have for the Jacobian

$$\begin{aligned} \frac{\partial(f_1, f_2)}{\partial(z_1, z_2)} &= \begin{vmatrix} 2z_1 - 2a & 2z_1 - a - b \\ -2z_2 + 2b & -2z_2 + a + b \end{vmatrix} \\ &= 2(a-b)(z_1 + z_2 - a - b) \\ &= 2(f_2 - f_1). \end{aligned}$$

Thus the Jacobian does not vanish when either f_1 or f_2 vanishes *alone*: nor, we may observe, does any first minor of the Jacobian vanish, even

if $f_1 = 0 = f_2$. The conditions of (35) are therefore not satisfied, but nevertheless the pair of equations (39) can be seen to have the eliminant

$$(a-b)(a^2-b^2+1) = 0.$$

We have thus obtained, in (33), conditions that are necessary but not sufficient for the existence of an eliminant, and, in (35), conditions that are sufficient but not necessary. It only remains to add that I do not know any way of bridging this gap and of finding exact conditions, i.e. conditions at once necessary and sufficient.

11. Elimination of more or fewer variables

The foregoing arguments can be extended to discuss the elimination of more than n variables between n equations: let us say $n+r$ variables between n equations. In analogy with (33) and (35) we can establish the following theorems which I give without proof.

(40) *If the $n+r$ variables $z_1,\ldots,z_{n+r}$ can be eliminated between the n equations $f_1 = 0, \ldots, f_n = 0$, then it is necessary that the set of Jacobians*

$$\frac{\partial(f_1,\ldots,f_n)}{\partial((z_1,\ldots,z_{n+r}))}$$

all vanish when $f_1,\ldots,f_n$ all vanish.

and

(41) *The $n+r$ variables $z_1,\ldots,z_{n+r}$ can be eliminated between the n equations $f_1 = 0, \ldots, f_n = 0$, if the set of Jacobians*

$$\frac{\partial(f_1,\ldots,f_n)}{\partial((z_1,\ldots,z_{n+r}))}$$

all vanish whenever $n-1$ only of the given equations are satisfied, say $f_2 = 0, \ldots, f_n = 0$, provided only that the minors independent of f_1 in the above Jacobians, i.e. the set of Jacobians

$$\frac{\partial(f_2,\ldots,f_n)}{\partial((z_1,\ldots,z_{n+r}))},$$

are not all zero when $f_2 = 0, \ldots, f_n = 0$ simultaneously.

Here again I cannot, of course, give an exact theorem bridging the gap between (40) and (41).

On the other hand we do not require conditions for the elimination of *less* than n variables between n equations: let us say the $n-r$ variables $z_1,\ldots,z_{n-r}$ between the n equations $f_1 = 0, \ldots, f_n = 0$. For either (i) we can eliminate $z_1,\ldots,z_{n-r}$ between the $n-r$ equations $f_1 = 0, \ldots, f_{n-r} = 0$, getting one eliminant of the given equations, and

obtaining, in similar fashion, other eliminants (not necessarily all independent) between every set of $n-r$ of the given equations; or else (ii) we can solve at least one set of $n-r$ of the equations in $z_1,\ldots,z_{n-r}$. The substitution of these values for the z's in the remaining r equations provides r eliminants.

We ought to suppose that the elements, other than $z_1,\ldots,z_{n-r}$, in the given equations are distinct, for otherwise the given equations may be redundant, i.e. no further condition is needed to make them consistent, as, for example, in the set of equations

$$z_1+z_2 = a+1, \qquad z_1^2+z_2^2 = a^2+1, \qquad z_1 z_2 = a. \tag{42}$$

In such an example it is better to begin with the more general set

$$z_1+z_2 = A, \qquad z_1^2+z_2^2 = B, \qquad z_1 z_2 = C, \tag{43}$$

which has the eliminant $A^2 = B+2C$, and then to say that in (42) the corresponding eliminant vanishes identically.

It may happen exceptionally in case (ii) that the implicit functions $z_1,\ldots,z_{n-r}$ have a domain that in the real field is restricted to a single point. The substitution of these unique values in the other r equations is still sufficient to furnish the eliminants. If in the extreme case the implicit functions z have no real domain, the given equations are inconsistent in the real field and we had best say so at once.

12. Elimination between functions

We may proceed from elimination between *equations* to elimination between *functions*. Given n functions of $z_1,\ldots,z_n$,

$$u_r(z_1,\ldots,z_n) \qquad (r = 1,\ldots,n), \tag{44}$$

we shall say that the variables z have been eliminated between the functions u, if we have found a function of the u's only

$$\psi(u_1,\ldots,u_n) \tag{45}$$

which is independent of the z's. This function ψ is thus defined by a relation of the form

$$\psi\{u_1(z_1,\ldots,z_n),\ldots,u_n(z_1,\ldots,z_n)\} = C,$$

where C is independent of $z_1,\ldots,z_n$. We may therefore absorb C into the function by writing

$$\phi(u_1,\ldots,u_n) = \psi(u_1,\ldots,u_n)-C,$$

so that now
$$\phi\{u_1(z_1,\ldots,z_n),\ldots,u_n(z_1,\ldots,z_n)\} = 0 \tag{46}$$

for every $z_1,\ldots,z_n$.

We may go further and investigate the conditions that the z's be

expressible as functions of the u's, i.e. that functions ϕ_r can be found such that

$$z_r = \phi_r(u_1,\ldots,u_n) \qquad (r = 1,\ldots,n).$$

We may, if we please, call this 'solving the functions' in terms of the z's.

This can be made part of our previous theory of solving and eliminating between *equations*, if we suppose that the n variables z are to be determined by the n equations

$$f_r(z_1,\ldots,z_n;u_1,\ldots,u_n) \equiv \phi_r(z_1,\ldots,z_n)-u_r = 0 \qquad (r = 1,\ldots,n). \quad (47)$$

The desired functional relation (46) now appears as the eliminant of the n equations (47).

The Jacobian
$$\frac{\partial(f_1,\ldots,f_n)}{\partial(z_1,\ldots,z_n)} = \frac{\partial(\phi_1,\ldots,\phi_n)}{\partial(z_1,\ldots,z_n)} \quad (48)$$

is now independent of the n parameters u_r and therefore, if it vanish at all, vanishes independently of the equations (44), since each of these contains just one parameter u_r which appears neither in the Jacobian nor elsewhere in the equations. The Jacobian, then, either vanishes identically or not at all. Similarly, if its first minors vanish, they too vanish identically.

Thus, by (33) and (35), the vanishing of the Jacobian (48) is both a necessary and a sufficient condition for the existence of the eliminant (46), unless every first minor of the Jacobian vanish identically. If every first minor vanish identically, the Jacobian itself vanishes identically; the identical vanishing of the Jacobian is therefore certainly a necessary condition for the existence of the eliminant (46). This, of course, can be proved directly by differentiating (46) with respect to every z_r in turn.

Conversely, we have so far proved that the identical vanishing of the Jacobian is a sufficient condition for the existence of the eliminant (46), unless every first minor vanishes identically. But, by repetition of the argument, identical vanishing of the first minor

$$\frac{\partial(\phi_2,\ldots,\phi_n)}{\partial(z_2,\ldots,z_n)} \quad (49)$$

is sufficient to establish the existence of an eliminant

$$\theta(u_2,\ldots,u_n) = 0, \quad (50)$$

unless every first minor of (49) vanish identically.

Proceeding in this way we at length arrive at an eliminant (or more probably a set of eliminants) unless every $(n-1)$th minor $\partial f_r/\partial z_s$ of the Jacobian vanish identically. In this extreme case every f is independent

of every z and we may say either that there are n eliminants, or, with more candour, that the given functions were merely fraudulent.

At any rate, the identical vanishing of the Jacobian is to be regarded as sufficient for the existence of the eliminant. Observing that, in virtue of the equations (47), we may write the Jacobian (48) as

$$\frac{\partial(u_1,\ldots,u_n)}{\partial(z_1,\ldots,z_n)},$$

we may sum up our conclusions in the theorem:

(51) *Given n functions of n variables $z_1,\ldots,z_n$*

$$u_r(z_1,\ldots,z_n) \qquad (r = 1,\ldots,n),$$

the identical vanishing of their Jacobian

$$\frac{\partial(u_1,\ldots,u_n)}{\partial(z_1,\ldots,z_n)}$$

is a necessary and sufficient condition for the existence of an eliminant

$$\phi(u_1,\ldots,u_n) = 0.$$

From (51) we deduce a theorem fundamental in the study of the linear first-order partial differential equation

$$X_1\frac{\partial u}{\partial x_1}+X_2\frac{\partial u}{\partial x_2}+\ldots+X_n\frac{\partial u}{\partial x_n} = 0,$$

where u is the unknown function and X_1, $X_2,\ldots,$ X_n given functions of the independent variables $x_1,\ldots,x_n$, namely:

(52) *If $u_1,\ldots,u_{n-1}$ are $n-1$ functionally distinct solutions of the partial differential equation*

$$X_1\frac{\partial u}{\partial x_1}+\ldots+X_n\frac{\partial u}{\partial x_n} = 0,$$

then the general solution is of the form

$$u = \phi(u_1,\ldots,u_{n-1}),$$

where ϕ is an arbitrary function.

For, if u be any solution, we have

$$X_1\frac{\partial u}{\partial x_1}+\ldots+X_n\frac{\partial u}{\partial x_n} = 0$$

and

$$X_1\frac{\partial u_1}{\partial x_1}+\ldots+X_n\frac{\partial u_1}{\partial x_n} = 0,$$

$$\cdots\cdots\cdots\cdots$$

$$X_1\frac{\partial u_{n-1}}{\partial x_1}+\ldots+X_n\frac{\partial u_{n-1}}{\partial x_n} = 0.$$

Elimination of $X_1,\dots, X_n$ gives

$$\frac{\partial(u, u_1,\dots, u_{n-1})}{\partial(x_1,\dots, x_n)} = 0,$$

i.e., by (51), $u, u_1,\dots, u_{n-1}$ are connected by a functional relation

$$f(u, u_1,\dots, u_{n-1}) = 0.$$

If u is absent from this relation, then $u_1,\dots, u_{n-1}$ are functionally connected, which we have excluded by hypothesis. We may therefore suppose it solved for u, and write

$$u = \phi(u_1,\dots, u_{n-1}),$$

which proves the theorem.

We may without difficulty extend the foregoing arguments to prove that

(53) *Given n functions of $n+r$ variables $z_1,\dots, z_{n+r}$*

$$u_s(z_1,\dots, z_{n+r}) \qquad (s = 1,\dots, n),$$

the identical vanishing of the set of Jacobians

$$\frac{\partial(u_1,\dots, u_n)}{\partial((z_1,\dots, z_{n+r}))}$$

is a necessary and sufficient condition for the existence of an eliminant

$$\phi(u_1,\dots, u_n) = 0.$$

It follows as a corollary that

(54) *Given n functions of n variables $z_1,\dots, z_n$,*

$$u_s(z_1,\dots, z_n) \qquad (s = 1,\dots, n),$$

a necessary and sufficient condition for the existence of r distinct eliminants

$$\phi_s(u_s, u_{r+1}, u_{r+2},\dots, u_n) = 0 \qquad (s = 1,\dots, r)$$

is that every $(r-1)$th minor of the Jacobian

$$\frac{\partial(u_1,\dots, u_n)}{\partial(z_1,\dots, z_n)}$$

vanish identically.

Finally, it is a mere restatement of (21) or of (51) to assert that

(55) *Given n functions of n variables $z_1,\dots, z_n$,*

$$u_r(z_1,\dots, z_n) \qquad (r = 1,\dots, n),$$

we can express the z's as functions of the u's, in the form

$$z_s = \phi_s(u_1,\dots, u_n) \qquad (s = 1,\dots, n),$$

in a region in which the Jacobian

$$\frac{\partial(u_1,\dots, u_n)}{\partial(z_1,\dots, z_n)}$$

does not vanish.

WORKED EXAMPLE

Obtain the envelope of the straight line

$$\theta^5+10A_2\theta^3+10A_3\theta^2+5\theta y+x = 0,$$

where θ is a variable parameter and A_2, A_3 constants.

In the several regions into which the plane is divided by this envelope, determine the number of branches of the implicit function $\theta(x,y)$ defined by the above quintic equation.

Obtain criteria for the number of real roots of the quintic equation.

Differentiation in θ gives for a typical point of the envelope

$$\left.\begin{aligned} x &= 4\theta^5+20A_2\theta^3+10A_3\theta^2 \\ y &= -\theta^4-6A_2\theta^2-4A_3\theta \end{aligned}\right\}. \tag{1}$$

We look first for cusps and nodes. At a cusp $dx = 0 = dy$, i.e.

$$20(\theta^4+3A_2\theta^2+A_3\theta)\,d\theta = 0 = -4(\theta^3+3A_2\theta+A_3)\,d\theta.$$

The parameters of the cusps are therefore given by the cubic equation

$$\theta^3+3A_2\theta+A_3 = 0. \tag{2}$$

The discriminant of this cubic is known to be $-(A_2^3+4A_3^2)$. Thus the cusps are all real, if and only if

$$A_2^3+4A_3^2 < 0. \tag{3}$$

It is, of course, incidentally necessary, but not sufficient, that $A_2 < 0$.

Again, a node corresponds to a pair of double roots of the quintic, which are real or imaginary according as the nodal tangents are real or imaginary, so that an isolated or conjugate point corresponds to a *double* pair of *conjugate* roots. Thus, if (x,y) is a node at which the two values of the parameter are given by

$$\theta^2-2\alpha\theta+\beta = 0, \tag{4}$$

we have an identity of the form

$$\theta^5+10A_2\theta^3+10A_3\theta^2+5y\theta+x \equiv (\theta+4\alpha)(\theta^2-2\alpha\theta+\beta)^2, \tag{5}$$

the residual factor on the right being chosen to secure the disappearance of the term in θ^4. Equating coefficients, we have

$$\beta-6\alpha^2 = 5A_2, \qquad 2\alpha\beta+8\alpha^3 = 5A_3, \tag{6}$$

$$\beta^2-16\alpha^2\beta = 5y, \qquad 4\alpha\beta^2 = x. \tag{7}$$

From (6) we get
$$4\alpha^3+2A_2\alpha-A_3 = 0 \tag{8}$$
and
$$\beta = 6\alpha^2+5A_2. \tag{9}$$

Substituting for (β) in (7) and reducing by means of (8), we have

$$\left.\begin{aligned} x &= 36A_3\alpha^2+16A_2^2\alpha+42A_2A_3 \\ y &= 2A_2\alpha^2-3A_3\alpha+5A_2^2 \end{aligned}\right\}. \tag{10}$$

The three values of α, real or imaginary, given by (8) determine, when substituted in (10), three real or imaginary nodes. The envelope is, from (1), of degree five and, from its defining equation, of class five. It is unicursal, since its coordinates are expressed rationally in terms of the parameter θ. It is therefore of deficiency zero. Plucker's equations then give, if δ, κ denote, as usual, the numbers of nodes and cusps,

$$2\delta+3\kappa = 15, \qquad \delta+\kappa = 6,$$

i.e. $\delta = 3$, $\kappa = 3$, the numbers of nodes and cusps that we have just found, which are therefore complete.

From (4), (9) we have for the pair of nodal parameters

$$\theta^2-2\alpha\theta+6\alpha^2+5A_2 = 0. \tag{11}$$

These are real, if and only if
$$\alpha^2+A_2 < 0. \tag{12}$$
But, from (8), we get by squaring and rearranging
$$(\alpha^2+A_2)(16\alpha^4+4A_2^2) = 4A_2^3+A_3^2.$$
Thus the tangents are real at a real node, if and only if
$$4A_2^3+A_3^2 < 0,$$
which is the inequality (3) already obtained as the condition for three real cusps.

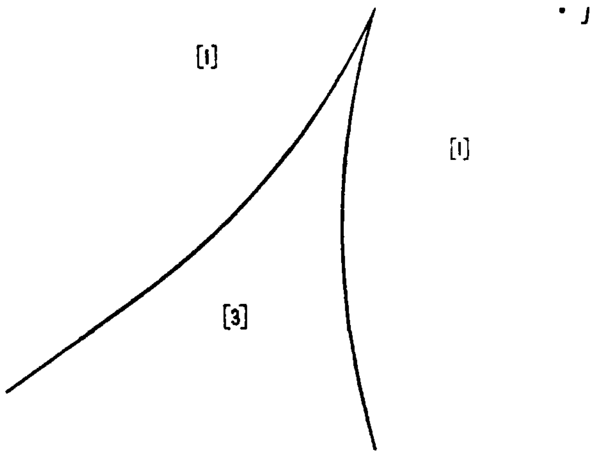

FIG 1 $(4A_2^3+A_3^2 > 0)$

Actually, the limiting case $4A_2^3+A_3^2 = 0$ gives coincident tangents at the node, as it also gives coincident cusps. As we shall see, in this case two cusps and a node have moved up to coincidence in a multiple point of the envelope. We further note that the roots of (8) will all be real, if
$$8A_2^3+27A_3^2 < 0, \tag{13}$$
an inequality necessitating (3) but not necessitated by it. Of these two inequalities it is (3) that is the important one, as we shall ultimately see.

The envelope can therefore have three possible types of figure.

(i) if $4A_2^3+A_3^2 > 0$, there is one real cusp and one isolated point J, but no other real singularity (Fig. 1);

(ii) if $4A_2^3+A_3^2 < 0 < 8A_2^3+27A_3^2$, there are three real cusps and one real node with real tangents, but no other real singularity (Fig. 2);

(iii) if $8A_2^3+27A_3^2 < 0$, there are three real cusps and three real nodes with real tangents (Fig. 3).

There are also the two intermediate limiting cases Figs. 1–2 and 2–3, where in Fig. 1–2 two cusps and a node of Fig. 2 coincide in a single multiple point P, and in Fig. 2–3 two nodes and a cusp of Fig. 3 coincide in a single point Q.

In drawing these curves we should remember that there are no finite points of inflexion, since the 'slope' of the tangent varies inversely as θ, and the tangent

therefore rotates steadily as the parameter θ moves steadily along the curve. At infinity the curve is to a first approximation

$$x = 4\theta^5, \qquad y = -\theta^4,$$

i.e.
$$(\tfrac{1}{4}x)^4 + y^5 = 0.$$

With those indications we can obtain the figures of the five curves given in the text.

In each of these figures the envelope (shown by the *full* line) divides the plane into a number of distinct regions ranging from two in Fig 1 to five in Fig 3.

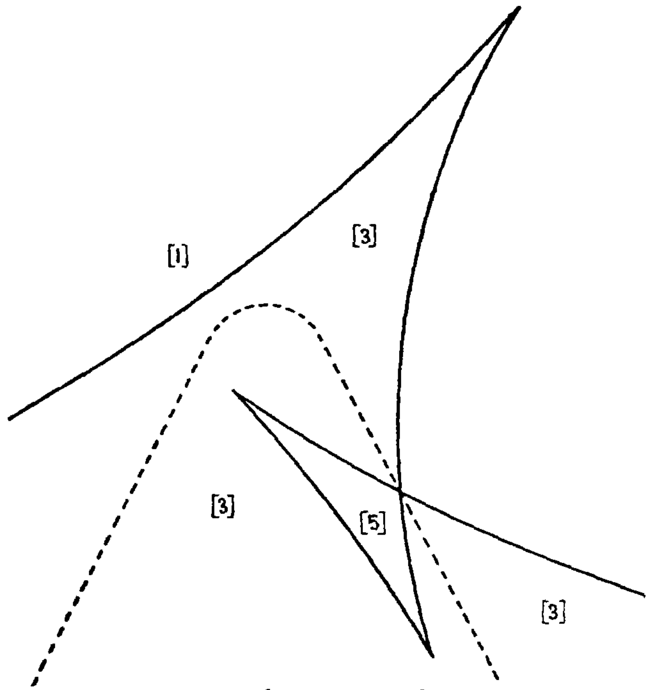

FIG 2 ($4A_2^3 + A_3^2 < 0 < 8A_2^3 + 27A_3^2$)

To determine the number of branches of the implicit function $\theta(x, y)$ in these several regions we observe that the number of real values of $\theta(x, y)$ at a given point (x, y) is the number of real tangents that can be drawn to the curve from that point. These we determine by inspection, i.e. visually. We are helped in doing this, if we regard the curve as divided up by consecutive cusps into a series of 'arcs'. To such an arc we can draw no real tangents from points on its concave side. From points on the convex side of the arc in the region bounded by the arc and the cuspidal tangents two real tangents can be drawn to the arc. If we leave this region by crossing one of the cuspidal tangents, one of the tangents to the arc becomes, of course, a tangent to the neighbouring arc terminated by the cusp. Thus one arc loses a tangent and the other gains one, the aggregate number of real tangents remaining unaltered. But, if we cross an arc by passing from its convex to its concave side, we lose two real tangents. Moreover, it is only by

crossing the envelope itself that we make any change in the total number of tangents to the envelope By working on these principles we at length assign to the implicit function $\theta(x, y)$ the numbers of real branches 1, 3, 5 as shown in brackets [ ] in the various regions of the several figures.

It should be remarked that in the limiting case shown by Fig. 1–2 the whole region [5] of Fig 2 has shrunk to the single point P. We must remember, too,

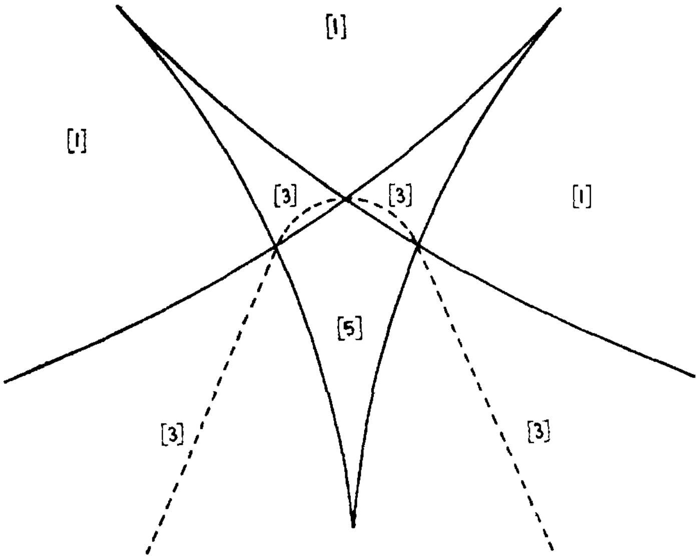

Fig 3 ($8A_2^3+27A_3^2 = 0$)

to ascribe the arcs of the envelope itself to the regions on the *convex* sides of these arcs, since the two tangents which become imaginary as we cross to the concave side of the arc remain real but coincident while we are actually on the arc itself. For this reason the point Q in Fig 2 3, although it is the limit of region [3] of Fig 3, must be assigned to the region [5] In point of fact the quintic represented by P has four roots $\sqrt[5]{(\frac{1}{2}A_3)}$ and one root $-4\sqrt[5]{(\frac{1}{2}A_3)}$, while the quintic represented by Q has three roots $2\sqrt[5]{A_3}$ and two roots $-3\sqrt[5]{A_3}$

We pass on now to establish criteria for the number of real and imaginary roots of the given quintic We should observe that, since A_2, A_3, x, y are independently at our disposal, any quintic can be reduced to the given form by removal of its second term, and that consequently criteria obtained for the given quintic are, in effect, criteria for the general quintic The coefficients A_2, A_3, x, y are seminvariants of the coefficients of the general quintic, and therefore any criteria that we obtain for the general quintic will be naturally expressed in terms of its seminvariants

Now, in the first place, we know from algebraic considerations that the sign of Δ the discriminant of the quintic distinguishes the respective cases of an even number and an odd number of conjugate pairs of imaginary roots. More precisely, if $\Delta > 0$, the quintic has 5 or 1 real roots, if $\Delta < 0$, the quintic has 3 real roots. Since $\Delta = 0$ is the condition for equal roots, it is, analytically, the

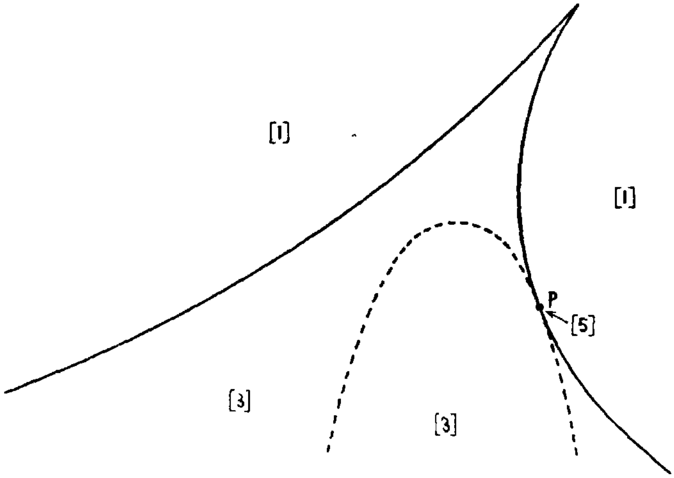

FIG. 1–2 $(4A_2^3+A_3^2=0)$

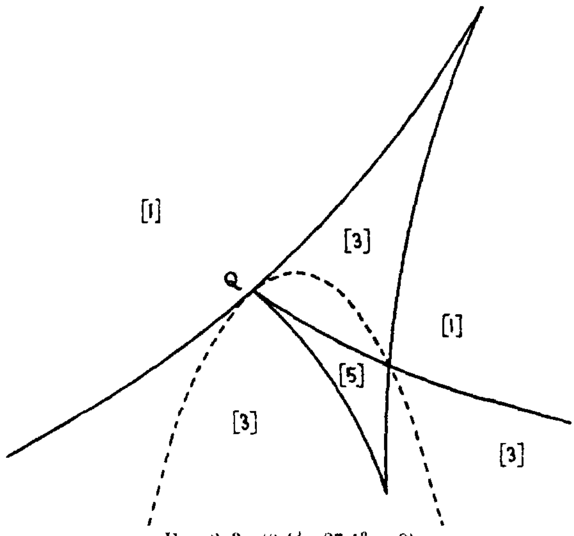

FIG. 2–3 $(8A_2^3+27A_3^2=0)$

equation to the envelope, and we find in the figures, as we should expect, that the envelope separates the regions [5] and [1] on the one side from the region [3] on the other. Thus in Fig. 1, in which there is no region [5], the sign of Δ is a sufficient criterion of the number of real roots. But in the other figures the sign of Δ does not suffice to distinguish the regions [5] and [1], which both correspond to positive Δ. We must therefore look for some further criterion to separate this pair of regions.

Now these two regions are mutually accessible only through the nodes of the

envelope. Thus any curve that passes through all the nodes of the envelope and does not cut the envelope elsewhere in real points will serve as a curve of separation of the regions [5] and [1].

Now the nodes are given above by (8), (10), i.e. by

$$x = 36A_3\alpha^2+16A_2^2\alpha+42A_2A_3,$$

$$y = 2A_2\alpha^2-3A_3\alpha+5A_2^2,$$

where $$4\alpha^3+2A_2\alpha-A_3 = 0.$$

Any α-eliminant of these three equations gives a curve through the nodes. The simplest such curves are the three conics

$$4y^2-A_3x-36A_2^2y+80A_2^4-54A_2A_3^2 = 0, \tag{14}$$

$$xy-4A_2^2x-18A_2A_3y-40A_2^3A_3+27A_3^3 = 0, \tag{15}$$

$$x^2-72A_2A_3x-(128A_2^4-108A_3^2)y+640A_2^5+432A_2^2A_3^2 = 0. \tag{16}$$

We must consider the 'residual' intersections of these curves with the envelope, i.e. the intersections other than the nodes. Now the elimination of α between (8), (11) gives

$$2\theta^6+22A_2\theta^4-14A_3\theta^3+68A_2^2\theta^2-72A_2A_3\theta+40A_2^3+27A_3^2 = 0 \tag{17}$$

as the sextic whose roots are the six parameters of the nodes. If we substitute for x, y from (1) into (14), (15), (16) respectively, we get (17) together with residual factors representing the residual intersections of the conics with the envelope. The residual factors are

$$2(\theta^2-A_2) = 0 \quad \text{for (14)},$$

$$-2\theta^3+A_3 = 0 \quad \text{for (15)};$$

$$8(\theta^4-A_2\theta^2-2A_3\theta-2A_2^2) = 0 \quad \text{for (16)}.$$

But in none of these three equations are all the roots imaginary under the single condition $4A_2^3+A_3^2<0$, which is the only condition we have to impose on the coefficients of the equations.

If, however, we combine (14), (16) into the hyperbola

$$\Theta \equiv x^2-16A_2y^2-76A_2A_3x-(272A_2^3-108A_3^2)y-24A_2^2(40A_2^3-27A_3^2) = 0, \tag{18}$$

we find that the residual intersections are given by

$$\theta^4-2A_3\theta+3A_2^2 = 0 \tag{19}$$

Now, by Newton's rule of signs, if this quartic have any real roots, they have the sign of A_3 and are at most two in number. But the discriminant of the quartic is a positive numerical multiple of

$$16A_2^6-A_3^4$$

and is therefore positive, if $$4A_2^3+A_3^2<0.$$

Hence, under the conditions of Figs. 2, 3, and 2 3, the quartic (19) has no real roots and $\Theta = 0$ is a suitable curve of separation. Actually only one branch of this hyperbola is relevant, namely that shown in broken line in Figs. 2, 3, and 2 3. The region [5] is on the concave side of this hyperbola, that is on the side $\Theta < 0$.

Remembering that $\Theta = 0$ passes through every node of $\Delta = 0$, including the isolated point of Fig. 1, we can now read from the figures the following necessary and sufficient criteria for real and imaginary roots of the quintic:

[5] *Five real*

$$\Delta > 0, \quad 4A_2^3+A_3^2 < 0, \quad \Theta \leqslant 0,$$
$$\Delta = 0, \quad 4A_2^3+A_3^2 = 0, \quad \Theta = 0$$

[3] *Three real, two imaginary*

$$\Delta < 0,$$
$$\Delta = 0, \quad 4A_2^3+A_3^2 > 0,$$
$$4A_2^3+A_3^2 = 0, \quad \Theta \neq 0,$$
$$4A_2^3+A_3^2 < 0, \quad \Theta = 0$$

[1] *One real, four imaginary*

$$\Delta > 0, \quad 4A_2^3+A_3^2 > 0,$$
$$4A_2^3+A_3^2 = 0, \quad \Theta > 0,$$
$$\Delta = 0, \quad 4A_2^3+A_3^2 = 0, \quad \Theta = 0$$

This last condition is necessary but not *sufficient*, since a further inequality is needed to distinguish the isolated point J from the two non real nodes Such a further inequality will not be simple, and, since the principle is clear, I shall not go to the pains of elaborating it for such an exceptional case

EXAMPLES XI

1 If z is defined implicitly as a function of $x_1, \ldots, x_n$ by the equation

$$z = a + x_1\phi_1(z) + \ldots + x_n\phi_n(z),$$

where the functions $\phi_r(z)$ are differentiable in z near $z = a$, show that the implicit function $z(x_1, \ldots, x_n)$ is bounded near the origin ($x_1 = 0, \ldots, x_n = 0$), and, if the derivatives $\phi_r'(z)$ are bounded near $z = a$, is also one valued and continuous

2 Show that the implicit function $z(x_1, x_2, x_3)$ defined by the equation

$$\frac{x_1^2}{z-a_1}+\frac{x_2^2}{z-a_2}+\frac{x_3^2}{z-a_3} = 1 \qquad (a_1 < a_2 < a_3)$$

has three branches z_1, z_2, z_3 such that

$$z_1 \to a_1, \quad z_2 \to a_2, \quad z_3 \to a_3,$$

and determine for what values of x_1, x_2, x_3 these branches coincide.

Show that, in a bounded region of the (x_1, x_2, x_3) space that excludes the planes $x_1 = 0$, $x_2 = 0$, $x_3 = 0$, the branches are distinct and continuous

3 Show that, if θ is a variable parameter, the envelope of the straight line

$$\theta^4 - 6\theta^2 - 4\theta y + 3x = 0$$

is a curve with cusps at $(-1, \pm 2)$ and a node at $(3, 0)$ Sketch the figure of this envelope and show how it divides up the plane into three regions in which the implicit function $\theta(x, y)$ defined by the above quartic equation has respectively 4, 2, 0 branches

Explain how the number of real roots of the quartic depends on the values of x, y, and prove, in particular, that all the roots of the quartic are real and distinct, if and only if its discriminant is positive and, in addition, $x < 3$

4. If θ is a variable parameter, sketch the figure of the envelope of the straight line

$$\theta^5 - 10\theta^3 - 5y\theta + 4x = 0,$$

showing that it has cusps at $(0, 0)$ and $(\pm 6\sqrt{3}, -9)$, and nodes at $(0, -5)$ and $(\pm 2\sqrt{2}, -4)$.

Indicate, for the various regions into which the (x, y) plane is divided by this envelope, the number of branches of the implicit function $\theta(x, y)$ defined by the above quintic equation.

Show that the necessary and sufficient condition that all the roots of the quintic be real and distinct is that its discriminant be positive and, in addition,

$$x^2 < 8y + 40$$

5. Sketch the envelope of the straight line

$$\theta^6 - 15\theta^4 + 60\theta^2 - 6\theta y + 5x = 0,$$

where θ is a variable parameter, showing that it has four cusps and six nodes at the respective points

$$(-18 \pm 10\sqrt{5}, \pm 8\sqrt{2}); \qquad (-10 \pm 2\sqrt{5}, 0), \quad (-11 \pm \sqrt{5}, \pm\sqrt{3})$$

Indicate, in the various regions of the (x, y) plane, the number of branches of the implicit function $\theta(x, y)$ defined by the above sextic equation

Show that the sextic has six real roots, if and only if

$$\text{(i)} \quad \Delta > 0, \qquad 57y^2 < (x - 2)(x^2 - 20x - 80), \qquad x - 2 < 0.$$

$$\Delta = 0, \qquad 57y^2 = (x - 2)(x^2 - 20x - 80), \qquad x - 2 < 0,$$

or equivalently

$$\text{(ii)} \quad \Delta > 0, \qquad 2x^2 + 40x + 160 - 3(5\sqrt{5} - 7)y^2 < 0,$$

$$\Delta = 0, \qquad 2x^2 + 40x + 160 - 3(5\sqrt{5} - 7)y^2 = 0$$

6. Sketch the envelope of the straight line

$$\theta^6 - 15\theta^4 + 75\theta^2 - 6\theta y + 5x = 0,$$

where θ is a variable parameter, showing that it has a pair of nodes at $(2, \pm 9)$, a pair of cusps at $(7, \pm 16)$, and a double cusp at $(-25, 0)$. Indicate, in the various regions of the (x, y) plane, the number of branches of the implicit function $\theta(x, y)$ defined by the above sextic equation

Show that the roots of the sextic are all real, only if it reduces to the limiting form

$$(\theta^2 - 5)^3 = 0.$$

7. (i) If (r, θ) and (r, θ, ϕ) are polar coordinates in two and three dimensions respectively, evaluate the Jacobians

$$\frac{\partial(r, \theta)}{\partial(x, y)}, \qquad \frac{\partial(r, \theta, \phi)}{\partial(x, y, z)},$$

where (x, y) and (x, y, z) are the corresponding rectangular cartesian coordinates.

(ii) If a, b, c, A, B, C are the sides and angles respectively of a spherical triangle, prove that

$$\frac{\partial(A, B, C)}{\partial(a, b, c)} = \frac{\sin A}{\sin a}.$$

8. (i) If $\phi(x, y, z, w)$, $\psi(x, y, z, w)$ are such as to satisfy three identities of the form

$$F(\phi, \psi, y, z) = 0, \qquad G(\phi, \psi, z, x) = 0, \qquad H(\phi, \psi, x, y) = 0,$$

show that ϕ, ψ are themselves functionally connected.

(ii) If $\phi(x, y, z, w)$, $\psi(x, y, z, w)$ are such as to satisfy two identities of the form

$$F(\phi, \psi, x, y) = 0, \qquad G(\phi, \psi, z, w) = 0,$$

show that ϕ, ψ are functionally connected, unless they each satisfy the differential equation

$$\frac{\partial\phi}{\partial x}\frac{\partial\phi}{\partial z}\frac{\partial^2\phi}{\partial y\partial w} + \frac{\partial\phi}{\partial y}\frac{\partial\phi}{\partial w}\frac{\partial^2\phi}{\partial x\partial z} = \frac{\partial\phi}{\partial x}\frac{\partial\phi}{\partial w}\frac{\partial^2\phi}{\partial y\partial z} + \frac{\partial\phi}{\partial y}\frac{\partial\phi}{\partial z}\frac{\partial^2\phi}{\partial x\partial w}.$$

9. (i) If $f(y, z)$, $f(z, x)$, $f(x, y)$ are connected by a functional relation, the form f being the same for all three functions, prove that $f(x, x)$ is independent of x.

(ii) Prove a similar result if

$$f(y, z)f(z, x), \qquad f(z, x)f(x, y), \qquad f(x, y)f(y, z)$$

are functionally related, and again, if

$$f(y, z)f(z, y), \qquad f(z, x)f(x, z), \qquad f(x, y)f(y, x)$$

are functionally related.

(iii) If $f(x, y, z)$, $f(y, z, x)$, $f(z, x, y)$ are functionally related, prove that $f(x, x, x)$ is independent of x.

If $f(x_1, x_2, x_3, x_4)$, $f(x_2, x_3, x_4, x_1)$, $f(x_3, x_4, x_1, x_2)$, $f(x_4, x_1, x_2, x_3)$ are functionally related, prove that either $f(x, x, x, x)$ or $f(x, -x, x, -x)$ is independent of x. [Non-real functional forms are to be deemed inadmissible.]

10. Given the four functions u, v, w, t in a field of three variables (x, y, z), prove that

(i) $$\begin{vmatrix} u_x & u_y & u_z & u \\ v_x & v_y & v_z & v \\ w_x & w_y & w_z & w \\ t_x & t_y & t_z & t \end{vmatrix} = 0, \qquad \text{(ii)} \begin{vmatrix} u_x & u_y & u_z & u \\ v_x & v_y & v_z & v \\ w_x & w_y & w_z & w \\ t_x & t_y & t_z & 0 \end{vmatrix} = 0,$$

(iii) $$\left\| \begin{matrix} u_x & u_y & u_z \\ v_x & v_y & v_z \end{matrix} \right\| = 0,$$

are respectively the conditions that

(i) u, v, w, t be connected by a homogeneous relation,

(ii) u, v, w, t be connected by a relation homogeneous in u, v, w,

(iii) u, v be functionally connected in the field of three variables.

11. Use the theory of § 12 to prove that the general solution of an ordinary differential equation of order n cannot contain more than n independent arbitrary constants.

12. Apply theorems (33) and (35) to the following pairs of equations, and determine whether the equations are soluble in z_1, z_2 or have an eliminant:

(i) $(z_1 + z_2 - a)^2 = 0, \qquad (z_1 - z_2 - b)^2 = 0;$

(ii) $(z_1 + z_2 - a)^2 = 0, \qquad (z_1^2 + z_1 z_2 - az_1 - b)^2 = 0;$

(iii) $z_1 + z_2 - a - b = 0, \qquad z_1^2 - z_2^2 - 2az_1 + 2bz_2 - 1 = 0.$

XII

CHANGE OF COORDINATES: JACOBIANS AND HESSIANS

1. Coordinate-systems

ANALYTICAL GEOMETRY consists essentially in the representation of points, lines, and other such geometrical elements by systems of coordinates, and the solution of a particular problem may be materially assisted by the apt choice of a coordinate-system. On the other hand, if the problem is really one of geometry, the choice of coordinate-system is extrinsic to the problem itself, which should remain unaltered in essentials for any other choice of coordinates. We are led therefore to consider the effect, on the analysis, of change of coordinate-system, of invariance under such change of coordinates, and so forth.

With the notable exception of polar coordinates, the various systems of coordinates normally used in Analytical Geometry are linearly related. In Differential Geometry, however, we recognize systems of coordinates of widest possible generality, restricting them, in fact, only by conditions of differentiability. Mathematical Physics, in so far as it is based on methods of Differential Geometry, introduces us to differential equations in such systems of coordinates, and we have therefore to consider the effect on differential equations and other differential expressions of general transformations of coordinates.

We have already, in chapter VI, discussed the behaviour of partial derivatives under certain changes of variable. Thus, in the simplest case, we considered three variables x, y, z connected by a relation $f(x, y, z) = 0$, and, in particular, expressed the partial derivatives of x (with respect to y, z) in terms of those of z (with respect to x, y). In this case there was an interchange of status between dependent and independent variable. In the systems we are to consider here the trans formation is essentially one of *independent* variables only: to emphasize this we have spoken, intentionally, of a transformation of *coordinates*. The dependent variable, a function of these coordinates, though changed, of course, in form by the transformation, remains essentially unchanged in meaning: in geometrical language we may regard it as a function of position, for example a potential.

2. Curvilinear coordinates

In a plane of cartesian coordinates (x, y), we define a new system of coordinates (u, v), when we define a pair of functions

$$u = u(x, y), \qquad v = v(x, y). \tag{1}$$

Suppose, conveniently, that these functions are one-valued. Then to a given pair of values x, y corresponds a definite pair of values u, v, which are the coordinates of the point (x, y) in the (u, v) system. The points of constant u-coordinate or constant v-coordinate trace out the *coordinate-curves*

$$u(x, y) = \text{constant}, \qquad v(x, y) = \text{constant},$$

which take the place of the rectangular network of the cartesian coordinate-curves

$$x = \text{constant}, \qquad y = \text{constant}.$$

For this reason these general coordinates are termed *curvilinear* coordinates. In similar fashion the homographic coordinates of analytical geometry may be distinguished as *rectilinear* coordinates.

To determine the cartesian coordinates of the point or points of intersection of a given pair of coordinate-curves $u =$ constant, $v =$ constant, we have to solve equations (1) in x, y, i.e. to determine the functions

$$x = x(u, v), \qquad y = y(u, v). \tag{2}$$

This is possible in a region in which the Jacobian

$$\frac{\partial(u, v)}{\partial(x, y)}$$

does not vanish. In a region throughout which this Jacobian vanishes u, v are connected by an identical relation $\phi(u, v) = 0$: between them, therefore, they dispose of but one degree of freedom and are incompetent to define a two-dimensional system. Accordingly, the generality of the functions (1) must be restricted by the condition that their Jacobian do not vanish. More generally, we may say that the region of validity of the substitution (1) is delimited by the curve

$$\frac{\partial(u, v)}{\partial(x, y)} = 0.$$

To this point we return later. In similar fashion the Jacobian of the functions (2) cannot vanish, since x, y are, by definition, essentially independent variables.

We may, if we please, regard equations (2) as an alternative definition of the curvilinear coordinates (u, v): in actual practice it is convenient to have both sets of equations (1), (2) available for easy passage between the two systems of coordinates. Equations (2) now define the coordinate-curves parametrically, for which reason they are sometimes called the *parametric* curves.

In three dimensions we can similarly define curvilinear or *surface* coordinates by either of the reciprocal substitutions

$$u = u(x,y,z), \qquad v = v(x,y,z), \qquad w = w(x,y,z); \tag{3}$$

$$x = x(u,v,w), \qquad y = y(u,v,w), \qquad z = z(u,v,w), \tag{4}$$

provided that the Jacobians of each set do not vanish; and so on for any n dimensions.

3. Orthogonal systems

The superiority of rectangular axes for metrical work in analytical geometry persists with curvilinear coordinates. Orthogonal coordinates, i.e. those in which the coordinate-curves and surfaces cut everywhere orthogonally, have natural advantages when we are dealing with differential forms based on the metrical quadratics x^2+y^2, $x^2+y^2+z^2$, etc., as, for example,

$$\frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2}+\frac{\partial^2 V}{\partial z^2}, \qquad (dx)^2+(dy)^2+(dz)^2.$$

In the plane the normal to the coordinate-curve $u(x,y) =$ constant has direction-ratios u_x, u_y, and an orthogonal system (u,v) is therefore defined by the single condition

$$u_x v_x + u_y v_y = 0. \tag{5}$$

Similarly, in three dimensions an orthogonal system (u,v,w) satisfies the three conditions

$$\left.\begin{aligned} v_x w_x + v_y w_y + v_z w_z &= 0 \\ w_x u_x + w_y u_y + w_z u_z &= 0 \\ u_x v_x + u_y v_y + u_z v_z &= 0 \end{aligned}\right\}, \tag{6}$$

and so on, $\frac{1}{2}n(n-1)$ conditions defining an orthogonal system in n dimensions.

If, on the other hand, the coordinates (u,v) are defined as parameters by (2), the tangent to $u =$ constant has direction-ratios x_v, y_v, and the orthogonal condition (5) takes the alternative form

$$x_u x_v + y_u y_v = 0. \tag{7}$$

Similarly, (4) leads to the set of conditions

$$\left.\begin{aligned} x_v x_w + y_v y_w + z_v z_w &= 0 \\ x_w x_u + y_w y_u + z_w z_u &= 0 \\ x_u x_v + y_u y_v + z_u z_v &= 0 \end{aligned}\right\}, \tag{8}$$

alternative to (6).

We discuss these conditions more carefully in a subsequent section. For mnemonic purposes we may note that the differential expressions

in (5), (6), (7), (8) are homogeneous in u, v, w severally, while, if these letters are suppressed, the fundamental metric x^2+y^2 or $x^2+y^2+z^2$ is formally reproduced.

It may be remarked that in a system of orthogonal coordinates the Jacobian vanishes, in general, only at isolated points. For, under the conditions (5),

$$\left\{\frac{\partial(u,v)}{\partial(x,y)}\right\}^2 = \begin{vmatrix} u_x^2+u_y^2 & 0 \\ 0 & v_x^2+v_y^2 \end{vmatrix},$$

which vanishes, in the real domain, only at the points

$$u_x = 0 = u_y, \qquad v_x = 0 = v_y;$$

and similarly in n dimensions. Of course, if u_x and u_y or v_x and v_y have a common factor, we shall get curves and not isolated points.

As examples of orthogonal coordinates in the plane we may mention

(i) polar coordinates:

$$r = \sqrt{(x^2+y^2)}, \qquad \theta = \tan^{-1}\left(\frac{y}{x}\right),$$

$$x = r\cos\theta, \qquad y = r\sin\theta,$$

in which the coordinate-curves are concentric circles and their diameters;

(ii) confocal coordinates:

$$u, v = \tfrac{1}{2}[x^2+y^2-a^2-b^2 \pm \sqrt{\{(x^2-y^2-a^2+b^2)^2+4x^2y^2\}}].$$

$$x = \sqrt{\frac{(a^2+u)(a^2+v)}{a^2-b^2}}, \qquad y = \sqrt{\frac{(b^2+u)(b^2+v)}{b^2-a^2}},$$

in which the coordinate-curves are the ellipses and hyperbolas of the confocal system

$$\frac{x^2}{a^2+\lambda}+\frac{y^2}{b^2+\lambda} = 1;$$

(iii) 'coaxal' coordinates:

$$u = \frac{x^2+y^2+a^2}{2x}, \qquad v = \frac{x^2+y^2-a^2}{2y},$$

$$x = \frac{u(v^2+a^2) \pm v\sqrt{\{(u^2-a^2)(v^2+a^2)\}}}{u^2+v^2},$$

$$y = \frac{v(u^2-a^2) \pm u\sqrt{\{(u^2-a^2)(v^2+a^2)\}}}{u^2+v^2},$$

in which the coordinate-curves are the circles of a coaxal system and of the corresponding orthogonal system.

4. Total differentials

We base our discussion of the transformation of coordinates on the formula of the total differential

$$Du = \frac{\partial u}{\partial x_1} Dx_1 + \ldots + \frac{\partial u}{\partial x_n} Dx_n, \tag{9}$$

and we therefore suppose that the functions defining the curvilinear coordinates are all fully differentiable.†

It is clear, of course, that this formula (9) is unique: in other words, if $\alpha_1,\ldots,\alpha_n$ are functions of $x_1,\ldots,x_n$ such that

$$Du = \alpha_1\, Dx_1 + \ldots + \alpha_n\, Dx_n$$

for all $Dx_1,\ldots, Dx_n$, then necessarily

$$\alpha_1 = \frac{\partial u}{\partial x_1}, \quad \ldots, \quad \alpha_n = \frac{\partial u}{\partial x_n}.$$

For we have only to take $Dx_2 = 0,\ldots, Dx_n = 0$ to get the first of these; and so on. That is to say, the *differential coefficients* $\partial u/\partial x_r$ are, in fact, sufficiently defined as being the *coefficients of the differentials* in the formula (9).

I find it simpler and more convincing, in what follows, to develop the analysis for the field of three coordinates only: the results appropriate to n dimensions can then be written down by analogy. In transforming, then, between the sets of coordinates (x,y,z) and (u,v,w) we have the total-differential formulae

$$\left.\begin{aligned} Du &= u_x\, Dx + u_y\, Dy + u_z\, Dz \\ Dv &= v_x\, Dx + v_y\, Dy + v_z\, Dz \\ Dw &= w_x\, Dx + w_y\, Dy + w_z\, Dz \end{aligned}\right\} \tag{10}$$

and

$$\left.\begin{aligned} Dx &= x_u\, Du + x_v\, Dv + x_w\, Dw \\ Dy &= y_u\, Du + y_v\, Dv + y_w\, Dw \\ Dz &= z_u\, Du + z_v\, Dv + z_w\, Dw \end{aligned}\right\}. \tag{11}$$

Solving (10) in Dx, say, we get

$$J\, Dx = \frac{\partial(v,w)}{\partial(y,z)} Du + \frac{\partial(w,u)}{\partial(y,z)} Dv + \frac{\partial(u,v)}{\partial(y,z)} Dw,$$

where

$$J \equiv \frac{\partial(u,v,w)}{\partial(x,y,z)}.$$

† In the sense of chapter VI § 2.

Since the formula of the total differential is unique, we can identify this relation with the first equation of (11), so getting

$$\frac{\partial x}{\partial u} = \frac{\partial(v,w)}{\partial(y,z)}\Big/\frac{\partial(u,v,w)}{\partial(x,y,z)}, \tag{12}$$

with other similar expressions. For mnemonic purposes we may observe that (12) is necessarily homogeneous in each of x, y, z, u, v, w. The corresponding formula in n dimensions is

$$\frac{\partial x_1}{\partial u_1} = \frac{\partial(u_2,\ldots,u_n)}{\partial(x_2,\ldots,x_n)}\Big/\frac{\partial(u_1,\ldots,u_n)}{\partial(x_1,\ldots,x_n)}. \tag{13}$$

It is just worth while to notice the corresponding formulae for the two-dimensional transformation from (x,y) to (u,v), namely

$$\frac{\partial x}{\partial u} = \frac{\partial v}{\partial y}\Big/J, \qquad \frac{\partial y}{\partial u} = -\frac{\partial v}{\partial x}\Big/J, \qquad \frac{\partial x}{\partial v} = -\frac{\partial u}{\partial y}\Big/J, \qquad \frac{\partial y}{\partial v} = \frac{\partial u}{\partial x}\Big/J, \tag{14}$$

where
$$J \equiv \frac{\partial(u,v)}{\partial(x,y)}.$$

There is a result due to Jacobi,† deducible from (12), which we shall need subsequently:

$$\frac{\partial}{\partial x}\left(J\frac{\partial x}{\partial u}\right)+\frac{\partial}{\partial y}\left(J\frac{\partial y}{\partial u}\right)+\frac{\partial}{\partial z}\left(J\frac{\partial z}{\partial u}\right) = 0. \tag{15}$$

The proof is immediate, for by (12) the left-hand side of (15) is

$$\frac{\partial}{\partial x}\left\{\frac{\partial(v,w)}{\partial(y,z)}\right\}+\frac{\partial}{\partial y}\left\{\frac{\partial(v,w)}{\partial(z,x)}\right\}+\frac{\partial}{\partial z}\left\{\frac{\partial(v,w)}{\partial(x,y)}\right\}.$$

Differentiation gives

$$\left\{\begin{vmatrix} v_{xy} & w_{xy} \\ v_z & w_z \end{vmatrix} - \begin{vmatrix} v_{zx} & w_{zx} \\ v_y & w_y \end{vmatrix}\right\} + \left\{\begin{vmatrix} v_{yz} & w_{yz} \\ v_x & w_x \end{vmatrix} - \begin{vmatrix} v_{xy} & w_{xy} \\ v_z & w_z \end{vmatrix}\right\} +$$
$$+\left\{\begin{vmatrix} v_{zx} & w_{zx} \\ v_y & w_y \end{vmatrix} - \begin{vmatrix} v_{yz} & w_{yz} \\ v_x & w_x \end{vmatrix}\right\},$$

in which the determinants cancel in pairs. Corresponding to (15) we have in n dimensions the formulae

$$\sum_{r=1}^{n}\frac{\partial}{\partial x_r}\left(J\frac{\partial x_r}{\partial u_s}\right) = 0 \qquad (s = 1,\ldots,n), \tag{16}$$

where
$$J \equiv \frac{\partial(u_1,\ldots,u_n)}{\partial(x_1,\ldots,x_n)}.$$

† *J. fur Math.* 27 (1844), 201–9; *Collected Works*, 4, 317.

5. An important analogy

From the general transformation of curvilinear coordinates we can derive an analogy with the linear transformation which is of far reaching importance. For this purpose we regard the point (x, y, z), and therefore the partial derivatives $\partial u/\partial x$, etc., evaluated at this point, as fixed, but we keep arbitrary the direction of differentiation D. This leaves us with Dx, Dy, Dz as the three ultimate 'coordinates', and Du, Dv, Dw as the three ultimate transformed coordinates. The equations of transformation are the equations of total differential (10), which, by a characteristic property of the total differential, are linear. These new differential coordinates accordingly transform by a linear transformation: they are, we may say, rectilinear coordinates, and we may assign to them the well-known properties of such coordinates.

The determinant of the transformation is the Jacobian J, and the condition $J \neq 0$ of linear independence of Du, Dv, Dw is the condition of functional independence of u, v, w. The conjugate or reversing substitution is (11), in which the coefficients are $\partial x/\partial u$, etc. The relation (12) then appears as the usual formula giving the coefficients of the reversing substitution in terms of those of the original substitution.

We may illustrate this principle in physical terms by taking time as the tacit variable of the differentiation D. The differentials Dx, Dy, Dz, Du, Dv, Dw become components of velocity: let us say, of a fluid in some hydrodynamical system. The linear equations are now those connecting the components of velocity of the fluid at a point of one system with the components of velocity of the fluid at the corresponding point in the other system. As we change from point to point, the coefficients of the linear transformation, and therefore the transformation itself, change also: the relative frame of reference of the two sets of component velocities or displacements $Dx,\ldots$, $Du\ldots$ modifies in scale and shape. But, so long as we are able to consider only displacements at a fixed point, the transformation retains the features of a linear transformation.

6. The transformation of orthogonal coordinates

The foregoing analogy retains its force, and the formulae themselves simplify in a transformation of orthogonal coordinates. Such coordinates, by definition, are subject to the set of conditions

$$\left.\begin{aligned} v_x w_x + v_y w_y + v_z w_z &= 0 \\ w_x u_x + w_y u_y + w_z u_z &= 0 \\ u_x v_x + u_y v_y + u_z v_z &= 0 \end{aligned}\right\}. \tag{17}$$

Define, in addition,

$$\left.\begin{aligned} u_x^2+u_y^2+u_z^2 &\equiv U^2 \\ v_x^2+v_y^2+v_z^2 &\equiv V^2 \\ w_x^2+w_y^2+w_z^2 &\equiv W^2 \end{aligned}\right\}. \tag{18}$$

These six equations show that

$$\left(\frac{u_x}{U},\frac{u_y}{U},\frac{u_z}{U}\right), \quad \left(\frac{v_x}{V},\frac{v_y}{V},\frac{v_z}{V}\right), \quad \left(\frac{w_x}{W},\frac{w_y}{W},\frac{w_z}{W}\right) \tag{19}$$

are the cosines of a triad of mutually perpendicular directions. Hence also

$$\left(\frac{u_x}{U},\frac{v_x}{V},\frac{w_x}{W}\right), \quad \left(\frac{u_y}{U},\frac{v_y}{V},\frac{w_y}{W}\right), \quad \left(\frac{u_z}{U},\frac{v_z}{V},\frac{w_z}{W}\right)$$

are the cosines of another such triad, and we have the six further equations

$$\left.\begin{aligned} \frac{u_y u_z}{U^2}+\frac{v_y v_z}{V^2}+\frac{w_y w_z}{W^2} &= 0 \\ \frac{u_z u_x}{U^2}+\frac{v_z v_x}{V^2}+\frac{w_z w_x}{W^2} &= 0 \\ \frac{u_x u_y}{U^2}+\frac{v_x v_y}{V^2}+\frac{w_x w_y}{W^2} &= 0 \end{aligned}\right\} \tag{20}$$

and

$$\left.\begin{aligned} \frac{u_x^2}{U^2}+\frac{v_x^2}{V^2}+\frac{w_x^2}{W^2} &= 1 \\ \frac{u_y^2}{U^2}+\frac{v_y^2}{V^2}+\frac{w_y^2}{W^2} &= 1 \\ \frac{u_z^2}{U^2}+\frac{v_z^2}{V^2}+\frac{w_z^2}{W^2} &= 1 \end{aligned}\right\}. \tag{21}$$

Write the total differential equations as

$$\left.\begin{aligned} \frac{Du}{U} &= \frac{u_x}{U}\,Dx+\frac{u_y}{U}\,Dy+\frac{u_z}{U}\,Dz \\ \frac{Dv}{V} &= \frac{v_x}{V}\,Dx+\frac{v_y}{V}\,Dy+\frac{v_z}{V}\,Dz \\ \frac{Dw}{W} &= \frac{w_x}{W}\,Dx+\frac{w_y}{W}\,Dy+\frac{w_z}{W}\,Dz \end{aligned}\right\}, \tag{22}$$

in which the coefficients on the right are, as we have said, the nine cosines of an orthogonal triad. This scheme of transformation of differentials has therefore the exact form of a transformation of

cartesian coordinates from one set of rectangular axes to another: let us say

$$\left.\begin{aligned} x' &= l_1 x+m_1 y+n_1 z \\ y' &= l_2 x+m_2 y+n_2 z \\ z' &= l_3 x+m_3 y+n_3 z \end{aligned}\right\}.$$

To solve these equations for x we multiply by l_1, l_2, l_3 and add. This, on a geometrical argument, gives the conjugate substitution

$$\left.\begin{aligned} x &= l_1 x'+l_2 y'+l_3 z' \\ y &= m_1 x'+m_2 y'+m_3 z' \\ z &= n_1 x'+n_2 y'+n_3 z' \end{aligned}\right\}.$$

Thus the transformation conjugate to (22) is

$$\left.\begin{aligned} Dx &= \frac{u_x}{U^2}\,Du+\frac{v_x}{V^2}\,Dv+\frac{w_x}{W^2}\,Dw \\ Dy &= \frac{u_y}{U^2}\,Du+\frac{v_y}{V^2}\,Dv+\frac{w_y}{W^2}\,Dw \\ Dz &= \frac{u_z}{U^2}\,Du+\frac{v_z}{V^2}\,Dv+\frac{w_z}{W^2}\,Dw \end{aligned}\right\}. \tag{23}$$

We can therefore replace (12) by the formula

$$x_u = u_x/U^2, \tag{24}$$

where $$U^2 = u_x^2+u_y^2+u_z^2.$$

Evidently $$x_u^2+y_u^2+z_u^2 = U^{-2}, \tag{25}$$

and we can give (24) the symmetrical form

$$\frac{x_u}{\sqrt{(x_u^2+y_u^2+z_u^2)}} = \frac{u_x}{\sqrt{(u_x^2+u_y^2+u_z^2)}}.$$

Again, (24) enables us to write the orthogonal conditions (17) in the form

$$x_v x_w+y_v y_w+z_v z_w = 0, \quad \text{etc.}, \tag{26}$$

already given in (8) above by a geometrical argument. We may regard the triads of equations represented by (25), (26) as being the conjugates of (18), (17).

For the Jacobian of the orthogonal system we have

$$\begin{vmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix}^2 = \begin{vmatrix} U^2 & 0 & 0 \\ 0 & V^2 & 0 \\ 0 & 0 & W^2 \end{vmatrix},$$

in virtue of (17), (18), and so

$$J = \pm UVW.$$

Since U, V, W have been defined only through their squares, their signs are still ambiguous: we may therefore sufficiently take

$$J = UVW. \tag{27}$$

This has the effect of securing that the orthogonal triad (19) is a positive or right-handed triad.

The defining properties of rectangular cartesian coordinates can be epitomized by stating the invariance of the quadratic form

$$x^2+y^2+z^2 = x'^2+y'^2+z'^2.$$

In similar fashion the defining properties of orthogonal curvilinear coordinates are all deducible from the identity of total differentials

$$(Dx)^2+(Dy)^2+(Dz)^2 = \frac{(Du)^2}{U^2}+\frac{(Dv)^2}{V^2}+\frac{(Dw)^2}{W^2},$$

which we may obtain either from (22) or from (23).

In n dimensions we have the $\frac{1}{2}n(n-1)$ orthogonal conditions

$$\sum_{r=1}^{n} \frac{\partial u_s}{\partial x_r}\frac{\partial u_t}{\partial x_r} = 0 \qquad (s,t = 1,\ldots,n;\ s \neq t) \tag{28}$$

and we further define

$$\sum_{r=1}^{n} \left(\frac{\partial u_s}{\partial x_r}\right)^2 = U_s^2 \qquad (s = 1,\ldots,n). \tag{29}$$

Corresponding to (20), (21) we have

$$\sum_{r=1}^{n} \frac{\partial u_r}{\partial x_s}\frac{\partial u_r}{\partial x_t}\bigg/ U_r^2 = 0 \qquad (s,t = 1,\ldots,n;\ s \neq t), \tag{30}$$

$$\sum_{r=1}^{n} \left(\frac{\partial u_r}{\partial x_s}\right)^2 \bigg/ U_r^2 = 1 \qquad (s = 1,\ldots,n). \tag{31}$$

Corresponding to (24), (25) we have

$$\frac{\partial x_r}{\partial u_s} = \frac{\partial u_s}{\partial x_r}\bigg/ U_s^2, \tag{32}$$

$$\sum_{r=1}^{n} \left(\frac{\partial x_r}{\partial u_s}\right)^2 = U_s^{-2}. \tag{33}$$

The orthogonal conditions may also be written as

$$\sum_{r=1}^{n} \frac{\partial x_r}{\partial u_s}\frac{\partial x_r}{\partial u_t} = 0 \qquad (s,t = 1,\ldots,n;\ s \neq t), \tag{34}$$

and the Jacobian is

$$\frac{\partial(u_1,\ldots,u_n)}{\partial(x_1,\ldots,x_n)} = \prod_{r=1}^{n} U_r. \tag{35}$$

7. Properties of Jacobians

The association of successive linear transformations gives rise, as w know, to simple relations between the determinants of the transform tions. Analogy indicates corresponding relations between the Jacobia of curvilinear transformations similarly associated. These relations ma be deduced by analogy from the corresponding formulae of the line transformation. There is, however, no difficulty, except a certain u handiness of notation, in proving them directly.

The formulae in question are as follows:

(36) *If the n functions*

$$\phi_r(u_1,\dots,u_n) \qquad (r=1,\dots,n)$$

are defined as functions of $x_1,\dots,x_n$ *through the relations*

$$u_s = u_s(x_1,\dots,x_n) \qquad (s=1,\dots,n),$$

then
$$\frac{\partial(\phi_1,\dots,\phi_n)}{\partial(x_1,\dots,x_n)} = \frac{\partial(\phi_1,\dots,\phi_n)}{\partial(u_1,\dots,u_n)}\frac{\partial(u_1,\dots,u_n)}{\partial(x_1,\dots,x_n)}.$$

In particular,

(37) *If* $u_1,\dots,u_n$ *are given as functions of* $x_1,\dots,x_n$ *and vice versa t relations*

$$x_r = x_r(u_1,\dots,u_n), \quad u_r = u_r(x_1,\dots,x_n) \qquad (r=1,\dots,n),$$

then
$$\frac{\partial(x_1,\dots,x_n)}{\partial(u_1,\dots,u_n)}\frac{\partial(u_1,\dots,u_n)}{\partial(x_1,\dots,x_n)} = 1.$$

More generally,

(38) *Given n functions*

$$\phi_r(u_1,\dots,u_{m+n}) \qquad (r=1,\dots,n)$$

of $m+n$ *variables* $u_1,\dots,u_{m+n}$, *which are themselves functions of n variabl* $x_1,\dots,x_n$, *then*

$$\frac{\partial(\phi_1,\dots,\phi_n)}{\partial(x_1,\dots,x_n)} = \sum \frac{\partial(\phi_1,\dots,\phi_n)}{\partial(u_\alpha,\dots,u_\omega)}\frac{\partial(u_\alpha,\dots,u_\omega)}{\partial(x_1,\dots,x_n)},$$

where the summation $\sum$ *is taken over all distinct sets* $u_\alpha,\dots,u_\omega$, *these bein any n of* $u_1,\dots,u_{m+n}$.

Lastly,

(39) *If* $u_1,\dots,u_n$ *are defined implicitly as functions of* $x_1,\dots,x_n$ *by relation*

$$\phi_r(u_1,\dots,u_n;x_1,\dots,x_n) \qquad (r=1,\dots,n),$$

then
$$\frac{\partial(u_1,\dots,u_n)}{\partial(x_1,\dots,x_n)} = (-)^n\frac{\partial(\phi_1,\dots,\phi_n)}{\partial(x_1,\dots,x_n)}\Big/\frac{\partial(\phi_1,\dots,\phi_n)}{\partial(u_1,\dots,u_n)},$$

provided that the denominator does not vanish.

For conciseness we use the notation $|a_{rs}|$ for the determinant in which a_{rs} is a typical element. Then, in (36),

$$\frac{\partial(\phi_1,\ldots,\phi_n)}{\partial(x_1,\ldots,x_n)} = \left|\frac{\partial\phi_r}{\partial x_s}\right|$$

$$= \left|\sum_{t=1}^{n} \frac{\partial\phi_r}{\partial u_t}\frac{\partial u_t}{\partial x_s}\right|, \quad \text{by the formula of the total differential,}$$

$$= \left|\frac{\partial\phi_r}{\partial u_t}\right| \times \left|\frac{\partial u_t}{\partial x_s}\right|, \quad \text{by the determinantal product-formula,}$$

$$= \frac{\partial(\phi_1,\ldots,\phi_n)}{\partial(u_1,\ldots,u_n)}\frac{\partial(u_1,\ldots,u_n)}{\partial(x_1,\ldots,x_n)}.$$

This establishes (36). We prove (37) as a corollary by taking

$$\phi_r(u_1,\ldots,u_n) \equiv x_r(u_1,\ldots,u_n) \qquad (r = 1,\ldots,n).$$

Then $$\frac{\partial(\phi_1,\ldots,\phi_n)}{\partial(x_1,\ldots,x_n)} = \frac{\partial(x_1,\ldots,x_n)}{\partial(x_1,\ldots,x_n)} = 1.$$

In (38) we have

$$\frac{\partial(\phi_1,\ldots,\phi_n)}{\partial(x_1,\ldots,x_n)} = \left|\frac{\partial\phi_r}{\partial x_s}\right| = \left|\sum_{t=1}^{m+n} \frac{\partial\phi_r}{\partial u_t}\frac{\partial u_t}{\partial x_s}\right|.$$

This last determinant is the formal product of the two *arrays*

$$\left\|\frac{\partial\phi_r}{\partial u_t}\right\| \quad \begin{pmatrix} r = 1,\ldots,n \\ t = 1,\ldots,m+n \end{pmatrix}, \qquad \left\|\frac{\partial u_t}{\partial x_s}\right\| \quad \begin{pmatrix} t = 1,\ldots,m+n \\ s = 1,\ldots,n \end{pmatrix}.$$

But, by the theory of determinants, this formal product is expressible as the sum of products

$$\sum \left|\frac{\partial\phi_r}{\partial u_t}\right| \times \left|\frac{\partial u_t}{\partial x_s}\right| \qquad \begin{pmatrix} r, s = 1,\ldots,n \\ t = \alpha,\ldots,\omega \end{pmatrix},$$

where $\sum$ is taken over all $\alpha,\ldots,\omega$, these being any n of $1,\ldots,m+n$. Hence, at length,

$$\frac{\partial(\phi_1,\ldots,\phi_n)}{\partial(x_1,\ldots,x_n)} = \sum \frac{\partial(\phi_1,\ldots,\phi_n)}{\partial(u_\alpha,\ldots,u_\omega)}\frac{\partial(u_\alpha,\ldots,u_\omega)}{\partial(x_1,\ldots,x_n)}.$$

In (39), partial differentiation of any $\phi_r(u_1,\ldots,u_n;x_1,\ldots,x_n) = 0$ with respect to any x_s gives

$$\frac{\partial\phi_r}{\partial x_s} + \sum_{t=1}^{n} \frac{\partial\phi_r}{\partial u_t}\frac{\partial u_t}{\partial x_s} = 0.$$

Thus
$$\begin{aligned}\frac{\partial(\phi_1,\ldots,\phi_n)}{\partial(x_1,\ldots,x_n)} &= \left|\frac{\partial\phi_r}{\partial x_s}\right| \\ &= \left|-\sum_{t=1}^{n}\frac{\partial\phi_r}{\partial u_t}\frac{\partial u_t}{\partial x_s}\right| \\ &= (-)^n\left|\frac{\partial\phi_r}{\partial u_t}\right|\times\left|\frac{\partial u_t}{\partial x_s}\right| \\ &= (-)^n\frac{\partial(\phi_1,\ldots,\phi_n)}{\partial(u_1,\ldots,u_n)}\frac{\partial(u_1,\ldots,u_n)}{\partial(x_1,\ldots,x_n)},\end{aligned}$$
and we have (39), provided, of course, that the denominator does not vanish.

When $n = 1$, the formulae (36) to (39) take the elementary forms, which are fundamental of their kind:

(i) $\dfrac{d\phi}{dx} = \dfrac{d\phi}{du}\dfrac{du}{dx}$;

(ii) $\dfrac{dx}{du}\dfrac{du}{dx} = 1$;

(iii) $\dfrac{d\phi}{dx} = \sum\dfrac{\partial\phi}{\partial u_r}\dfrac{\partial u_r}{\partial x}$;

(iv) $\dfrac{du}{dx} = -\dfrac{\partial\phi}{\partial x}\Big/\dfrac{\partial\phi}{\partial u}$, if $\phi(x,u) = 0$.

The analogy of the Jacobian formulae with the elementary formulae is emphasized by the form of the Jacobian notation.

The fundamental theorem that, if $\dfrac{\partial(u_1,\ldots,u_n)}{\partial(x_1,\ldots,x_n)}$ vanish identically, then $u_1,\ldots,u_n$ are connected by a functional relation $\phi(u_1,\ldots,u_n) = 0$ has for its elementary analogue the proposition

(v) if $\dfrac{du}{dx} = 0$, then u is constant, i.e. some $\phi(u) = 0$.

We may employ (39) to give for this theorem the following alternative proof, which is more immediate though less fundamental than that of chapter XI § 12 (51).

Between the n given relations
$$u_r = u_r(x_1,\ldots,x_n) \qquad (r = 1,\ldots,n)$$
we can eliminate the $n-1$ arguments $x_2,\ldots,x_n$, getting the relation
$$\phi_1(u_1,\ldots,u_n;x_1) = 0.$$
By similar elimination of $x_2,\ldots,x_n$ respectively, we obtain in succession the $n-1$ other relations
$$\phi_r(u_1,\ldots,u_n;x_r) = 0 \qquad (r = 2,\ldots,n).$$

We may look on these n relations as defining the u's implicitly and apply (39). Then the given condition

$$\frac{\partial(u_1,\dots,u_n)}{\partial(x_1,\dots,x_n)} = 0 \quad \text{gives} \quad \frac{\partial(\phi_1,\dots,\phi_n)}{\partial(x_1,\dots,x_n)} = 0,$$

i.e.

$$\prod_{r=1}^{n} \frac{\partial\phi_r}{\partial x_r} = 0,$$

since ϕ_r involves no x except x_r, and the determinant consequently reduces to its principal diagonal.

One factor, at least, must vanish, say

$$\frac{\partial\phi_r}{\partial x_r} = 0.$$

This gives us that $\phi_r(u_1,\dots,u_n;x_r)$ is independent of x_r. Hence $\phi_r = 0$ is actually

$$\phi_r(u_1,\dots,u_n) = 0,$$

an identity connecting the n functions $u_1, \dots, u_n$.

8. Second-order derivatives

We have so far dealt only with the transformation of partial derivatives of the *first* order. Formulae, analogous to (12), (13), but, of course, less simple in character, can be obtained for the transformation of second derivatives. As before I discuss in detail only the field of three independent variables (x, y, z) or (u, v, w).

Now a single differentiation of u in an arbitrary parameter gave us equations (10) of the form

$$Du = u_x\, Dx + u_y\, Dy + u_z\, Dz, \tag{40}$$

which, solved for Dx, Dy, Dz, gave equations of the form

$$J\, Dx = \frac{\partial(v,w)}{\partial(y,z)}\, Du + \frac{\partial(w,u)}{\partial(y,z)}\, Dv + \frac{\partial(u,v)}{\partial(y,z)}\, Dw, \tag{41}$$

equivalent to

$$Dx = x_u\, Du + x_v\, Dv + x_w\, Dw, \tag{42}$$

where

$$J = \frac{\partial(u,v,w)}{\partial(x,y,z)}.$$

A second differentiation in the same arbitrary parameter gives

$$D^2u = u_x\, D^2x + u_y\, D^2y + u_z\, D^2z + \\ + \sum_{x,y,z} \{u_{xx}\,(Dx)^2 + 2u_{yz}\, DyDz\}, \tag{43}$$

with corresponding equations for D^2v, D^2w. Now $D^2u,\dots$, $D^2x,\dots$ enter these equations exactly as $Du,\dots$, $Dx,\dots$ enter equations (40). We may

therefore solve for D^2x, getting, as in (41),

$$J\, D^2x = \sum_{u,v,w} \frac{\partial(v,w)}{\partial(y,z)} D^2u - \sum_{u,v,w} \frac{\partial(v,w)}{\partial(y,z)} \sum_{x,y,z} \{u_{xx}\,(Dx)^2 + 2u_{yz}\, DyDz\},$$

or, as in (42),

$$D^2x = x_u\, D^2u + x_v\, D^2v + x_w\, D^2w - \\ - \frac{1}{J} \sum_{u,v,w} \frac{\partial(v,w)}{\partial(y,z)} \sum_{x,y,z} \{u_{xx}\,(Dx)^2 + 2u_{yz}\, DyDz\}. \qquad (44)$$

But, as in (43), we have

$$D^2x = x_u\, D^2u + x_v\, D^2v + x_w\, D^2w + \\ + \sum_{u,v,w} \{x_{uu}\,(Du)^2 + 2x_{vw}\, DvDw\}. \qquad (45)$$

Comparison of (44), (45) shows that x_{uu} and x_{vu} are respectively the coefficients of $(Du)^2$ and of $2DvDw$ in

$$-\frac{1}{J} \sum_{u,v,u} \frac{\partial(v,w)}{\partial(y,z)} \sum_{x,y,z} \{u_{xx}\,(Dx)^2 + 2u_{yz}\, DyDz\}.$$

Substitution from (41) for Dx, Dy, Dz in terms of Du, Dv, Dw therefore gives

$$-J^3 \frac{\partial^2 x}{\partial u^2} = \sum_{u,v,w} \frac{\partial(v,w)}{\partial(y,z)} \sum_{x,y,z} \left\{u_{xx}\left[\frac{\partial(v,w)}{\partial(y,z)}\right]^2 + 2u_{yz} \frac{\partial(v,w)}{\partial(z,x)} \frac{\partial(v,w)}{\partial(x,y)}\right\}, \quad (46)$$

$$-J^3 \frac{\partial^2 x}{\partial v \partial w} = \sum_{u,v,u} \frac{\partial(v,w)}{\partial(y,z)} \sum_{x,y,z} \left\{u_{xx} \frac{\partial(w,u)}{\partial(y,z)} \frac{\partial(u,v)}{\partial(y,z)} + \right. \\ \left. + u_{yz}\left[\frac{\partial(w,u)}{\partial(z,x)} \frac{\partial(u,v)}{\partial(x,y)} + \frac{\partial(u,v)}{\partial(z,x)} \frac{\partial(w,u)}{\partial(x,y)}\right]\right\}. \qquad (47)$$

We can rewrite these results rather more concisely as

$$-J^3 \frac{\partial^2 x}{\partial u^2} = \sum_{u,v,u} \frac{\partial(v,w)}{\partial(y,z)} \begin{vmatrix} u_{xx} & u_{xy} & u_{xz} & v_x & w_x \\ u_{xy} & u_{yy} & u_{yz} & v_y & w_y \\ u_{xz} & u_{yz} & u_{zz} & v_z & w_z \\ v_x & v_y & v_z & 0 & 0 \\ w_x & w_y & w_z & 0 & 0 \end{vmatrix}, \qquad (48)$$

$$-J^3 \frac{\partial^2 x}{\partial v \partial w} = \sum_{u,v,w} \frac{\partial(v,w)}{\partial(y,z)} \begin{vmatrix} u_{xx} & u_{xy} & u_{xz} & u_x & v_x \\ u_{xy} & u_{yy} & u_{yz} & u_y & v_y \\ u_{xz} & u_{yz} & u_{zz} & u_z & v_z \\ w_x & w_y & w_z & 0 & 0 \\ u_x & u_y & u_z & 0 & 0 \end{vmatrix}. \qquad (49)$$

We effect a further simplification, if we write

$$(\alpha\beta, y, z) \equiv \begin{vmatrix} u_{\alpha\beta} & v_{\alpha\beta} & w_{\alpha\beta} \\ u_y & v_y & w_y \\ u_z & v_z & w_z \end{vmatrix},$$

where α, β are any of x, y, z and repetitions are allowed. Then we have

$$-J^3\frac{\partial^2 x}{\partial u^2} = \begin{vmatrix} (xx,y,z) & (xy,y,z) & (xz,y,z) & v_x & w_x \\ (xy,y,z) & (yy,y,z) & (yz,y,z) & v_y & w_y \\ (xz,y,z) & (yz,y,z) & (zz,y,z) & v_z & w_z \\ v_x & v_y & v_z & 0 & 0 \\ w_x & w_y & w_z & 0 & 0 \end{vmatrix}, \tag{50}$$

$$-J^3\frac{\partial^2 x}{\partial v\partial w} = \begin{vmatrix} (xx,y,z) & (xy,y,z) & (xz,y,z) & u_x & v_x \\ (xy,y,z) & (yy,y,z) & (yz,y,z) & u_y & v_y \\ (xz,y,z) & (yz,y,z) & (zz,y,z) & u_z & v_z \\ w_x & w_y & w_z & 0 & 0 \\ u_x & u_y & u_z & 0 & 0 \end{vmatrix}. \tag{51}$$

These formulae simplify, as we might expect, for orthogonal coordinates. We begin now with the formula (24)

$$x_u = u_x U^{-2},$$

where

$$U^2 = u_x^2+u_y^2+u_z^2.$$

Differentiation in an arbitrary parameter gives

$$\begin{aligned} x_{uu}\,Du + x_{uv}\,Dv + x_{uw}\,Dw &= U^{-2}(u_{xx}\,Dx + u_{xy}\,Dy + u_{xz}\,Dz) - \\ &\quad -2u_x U^{-3}(U_x\,Dx + U_y\,Dy + U_z\,Dz). \end{aligned} \tag{52}$$

But

$$\begin{aligned} Dx &= x_u\,Du + x_v\,Dv + x_w\,Dw \\ &= u_x U^{-2}\,Du + v_x V^{-2}\,Dv + w_x W^{-2}\,Dw, \end{aligned}$$

where V^2, W^2 are defined similarly to U^2. Substituting this, and the corresponding formulae for Dy, Dz, in (52), we obtain from the coefficient of Du

$$x_{uu} = U^{-4}(u_{xx}u_x+u_{xy}u_y+u_{xz}u_z) - 2u_x U^{-5}(u_x U_x+u_y U_y+u_z U_z).$$

But, by differentiation of the formula for U^2,

$$UU_x = u_{xx}u_x+u_{xy}u_y+u_{xz}u_z,$$

and we may therefore write

$$\begin{aligned} \frac{\partial^2 x}{\partial u^2} &= U^{-5}\{U^2U_x - 2u_x(u_x U_x+u_y U_y+u_z U_z)\} \\ &= -U^{-5}\{(u_x^2-u_y^2-u_z^2)U_x + 2u_x u_y U_y + 2u_x u_z U_z\}. \end{aligned} \tag{53}$$

We may also note that

$$u_x U_x + u_y U_y + u_z U_z = U^{-1} \sum_{x,y,z} (u_{xx} u_x^2 + 2u_{yz} u_y u_z).$$

Again, from the coefficient of Dv in (52) we have

$$x_{uv} = U^{-2}V^{-2}(u_{xx} v_x + u_{xy} v_y + u_{xz} v_z) - \\ -2u_x U^{-3}V^{-2}(v_x U_x + v_y U_y + v_z U_z).$$

But from symmetry in u, v we must likewise have

$$x_{uv} = U^{-2}V^{-2}(v_{xx} u_x + v_{xy} u_y + v_{xz} u_z) - \\ -2v_x U^{-2}V^{-3}(u_x V_x + u_y V_y + u_z V_z).$$

Now, in virtue of the orthogonal property,

$$0 = \partial_x(u_x v_x + u_y v_y + u_z v_z) \\ = u_{xx} v_x + u_{xy} v_y + u_{xz} v_z + u_x v_{xx} + u_y v_{xy} + u_z v_{xz}.$$

Thus, if we add the two expressions for x_{uv}, the terms in $U^{-2}V^{-2}$ disappear, and we are left with

$$\frac{\partial^2 x}{\partial u \partial v} = -u_x U^{-3}V^{-2}(v_x U_x + v_y U_y + v_z U_z) - \\ -v_x U^{-2}V^{-3}(u_x V_x + u_y V_y + u_z V_z). \tag{54}$$

We may also observe that

$$v_x U_x + v_y U_y + v_z U_z = U^{-1} \sum_{x,y,z} \{u_{xx} v_x u_x + u_{yz}(u_y v_z + u_z v_y)\}, \text{ etc.} \tag{55}$$

The analogous formulae for derivatives of the third and higher orders are even less pleasing.

9. The transformation of differential operators

The transformation of the linear operator is immediate. In our usual three-dimensional field of coordinates (x, y, z) and (u, v, w) we have

$$\partial_x = u_x \partial_u + v_x \partial_v + w_x \partial_w, \quad \text{etc.}, \tag{56}$$

and conversely $\quad \partial_u = x_u \partial_x + y_u \partial_y + z_u \partial_z, \quad$ etc.,

i.e., in virtue of (12),

$$J\frac{\partial}{\partial u} = \frac{\partial(v, w)}{\partial(y, z)} \frac{\partial}{\partial x} + \frac{\partial(v, w)}{\partial(z, x)} \frac{\partial}{\partial y} + \frac{\partial(v, w)}{\partial(x, y)} \frac{\partial}{\partial z}. \tag{57}$$

If we compare the associated transformations of differentials and of differential operators, namely,

$$\left.\begin{aligned} Du &= u_x\,Dx + u_y\,Dy + u_z\,Dz \\ Dv &= v_x\,Dx + v_y\,Dy + v_z\,Dz \\ Dw &= w_x\,Dx + w_y\,Dy + w_z\,Dz \end{aligned}\right\}$$

and
$$\left.\begin{aligned}\partial_x &= u_x\partial_u+v_x\partial_v+w_x\partial_w\\ \partial_y &= u_y\partial_u+v_y\partial_v+w_y\partial_w\\ \partial_z &= u_z\partial_u+v_z\partial_v+w_z\partial_w\end{aligned}\right\},$$
we observe

(i) that the two transformations proceed in contrary senses: one from (u,v,w) to (x,y,z), the other from (x,y,z) to (u,v,w);

(ii) that the matrices, or schemes of coefficients, of the two transformations, namely
$$\begin{pmatrix} u_x & u_y & u_z\\ v_x & v_y & v_z\\ w_x & w_y & w_z\end{pmatrix} \quad\text{and}\quad \begin{pmatrix} u_x & v_x & w_x\\ u_y & v_y & w_y\\ u_z & v_z & w_z\end{pmatrix}$$
differ only in the interchange of rows and columns. Two linear transformations so related are said to be *contragredient* or *dual*, and the relationship has importance in invariant-theory. Here it may be noted as helping us to remember the formulae of transformation for linear differential operators.

The theory of transforming the quadratic operator we may discuss in terms of Laplace's operator
$$\nabla^2 = \partial_x^2+\partial_y^2+\partial_z^2$$
as being sufficiently typical and of classic interest. Define, then,
$$\begin{aligned}A,B,C &= \sum u_x^2,\quad \sum v_x^2,\quad \sum w_x^2,\\ F,G,H &= \sum v_x w_x,\quad \sum w_x u_x,\quad \sum u_x v_x,\end{aligned}$$
where $\sum$, in this section, always denotes summation over x, y, z. Now, if we square both sides of (56) or form the product of two such operators, we get operators quadratic in ∂_u, ∂_v, ∂_w together with operators linear in them. The linear terms arise where operators of one factor have operated on the *coefficients* of the other factor: the quadratic terms arise where the operators operate only on the operand, the coefficients being unaffected. The set of *quadratic* operators is formed, then, by mere algebraic combination of the coefficients. Thus from (56) and its corresponding equations we get
$$\nabla^2 = A\partial_u^2+B\partial_v^2+C\partial_w^2+2F\partial_v\partial_w+2G\partial_w\partial_u+2H\partial_u\partial_v+ \\ +\alpha\partial_u+\beta\partial_v+\gamma\partial_w, \quad (58)$$
where α, β, γ have still to be determined. Operation with (58) on u itself gives at once
$$\alpha = \nabla^2 u,$$
and so for β, γ.

To calculate $\nabla^2 u$ regard the three equations

$$\left.\begin{aligned} u_x^2+u_y^2+u_z^2 &= A \\ u_x v_x+u_y v_y+u_z v_z &= H \\ u_x w_x+u_y w_y+u_z w_z &= G \end{aligned}\right\}$$

as linear equations in u_x, u_y, u_z. They give

$$u_x = \frac{A}{J}\frac{\partial(v,w)}{\partial(y,z)}+\frac{H}{J}\frac{\partial(w,u)}{\partial(y,z)}+\frac{G}{J}\frac{\partial(u,v)}{\partial(y,z)}, \quad \text{etc.},$$

i.e., by (12),

$$u_x = \frac{A}{J}\left(J\frac{\partial x}{\partial u}\right)+\frac{H}{J}\left(J\frac{\partial x}{\partial v}\right)+\frac{G}{J}\left(J\frac{\partial x}{\partial w}\right), \quad \text{etc.}$$

Differentiating these equations in x, in y, and in z respectively and adding, we get

$$\nabla^2 u = \frac{A}{J}\sum\frac{\partial}{\partial x}\left(J\frac{\partial x}{\partial u}\right)+\frac{H}{J}\sum\frac{\partial}{\partial x}\left(J\frac{\partial x}{\partial v}\right)+\frac{G}{J}\sum\frac{\partial}{\partial x}\left(J\frac{\partial x}{\partial w}\right)+$$
$$+J\sum\frac{\partial x}{\partial u}\frac{\partial}{\partial x}\left(\frac{A}{J}\right)+J\sum\frac{\partial x}{\partial v}\frac{\partial}{\partial x}\left(\frac{H}{J}\right)+J\sum\frac{\partial x}{\partial w}\frac{\partial}{\partial x}\left(\frac{G}{J}\right).$$

By (15) above, the coefficients of A/J, H/J, G/J vanish and we are left with

$$\nabla^2 u = J\left\{\frac{\partial}{\partial u}\left(\frac{A}{J}\right)+\frac{\partial}{\partial v}\left(\frac{H}{J}\right)+\frac{\partial}{\partial w}\left(\frac{G}{J}\right)\right\}.$$

Substituting this for α, and similar expressions for β,γ, in (58), we have

$$\frac{1}{J}\nabla^2 = \frac{\partial}{\partial u}\left(\frac{A}{J}\frac{\partial}{\partial u}+\frac{H}{J}\frac{\partial}{\partial v}+\frac{G}{J}\frac{\partial}{\partial w}\right)+\frac{\partial}{\partial v}\left(\frac{H}{J}\frac{\partial}{\partial u}+\frac{B}{J}\frac{\partial}{\partial v}+\frac{F}{J}\frac{\partial}{\partial w}\right)+$$
$$+\frac{\partial}{\partial w}\left(\frac{G}{J}\frac{\partial}{\partial u}+\frac{F}{J}\frac{\partial}{\partial v}+\frac{C}{J}\frac{\partial}{\partial w}\right). \quad (59)$$

For the general homogeneous quadratic operator in x, y, z,

$$(a,b,c,f,g,h)(\partial_x,\partial_y,\partial_z)^2,$$

where a, b, c, f, g, h are constants, a corresponding formula holds. We have only to make the following interpretations on the right of (59):

$$A = (a,b,c,f,g,h)(u_x,u_y,u_z)^2,$$
$$F = (a,b,c,f,g,h)(v_x,v_y,v_z)(w_x,w_y,w_z),\dagger$$

and similarly for B, C, G, H. The generalization to many dimensions is straightforward.

If, in the transformation of ∇^2, the coordinates are orthogonal, then

† I use this to denote the 'polar' form

$$\sum a v_x w_x + \sum f(v_y w_z+v_z w_y).$$

F, G, H (in their old senses) vanish and we write U^2, V^2, W^2 for A, B, C; moreover, J is UVW. Thus, in orthogonal coordinates, we have

$$\nabla^2 = UVW\left\{\frac{\partial}{\partial u}\left(\frac{U}{VW}\frac{\partial}{\partial u}\right)+\frac{\partial}{\partial v}\left(\frac{V}{WU}\frac{\partial}{\partial v}\right)+\frac{\partial}{\partial w}\left(\frac{W}{UV}\frac{\partial}{\partial w}\right)\right\}. \tag{60}$$

In n dimensions we have similarly, in orthogonal coordinates,

$$\sum_r \frac{\partial^2}{\partial x_r^2} = J\sum_s \left\{\frac{\partial}{\partial u_s}\left(\frac{U_s^2}{J}\frac{\partial}{\partial u_s}\right)\right\}, \tag{61}$$

where $$U_s^2 = \sum_r \left(\frac{\partial u_s}{\partial x_r}\right)^2, \qquad J = \frac{\partial(u_1,\dots,u_n)}{\partial(x_1,\dots,x_n)}.$$

As an illustration of (60) consider ∇^2 in spherical-polar coordinates

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta.$$

Then $$x_r^2+y_r^2+z_r^2 = 1,$$
$$x_\theta^2+y_\theta^2+z_\theta^2 = r^2,$$
$$x_\phi^2+y_\phi^2+z_\phi^2 = r^2\sin^2\theta.$$

If we take $u, v, w \equiv r, \theta, \phi$, these give

$$1/U,\ 1/V,\ 1/W = 1,\ r,\ r\sin\theta,$$

and so

$$\nabla^2 = \frac{1}{r^2\sin\theta}\left\{\frac{\partial}{\partial r}\left(r^2\sin\theta\frac{\partial}{\partial r}\right)+\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{\partial}{\partial\phi}\left(\operatorname{cosec}\theta\frac{\partial}{\partial\phi}\right)\right\},$$

i.e. $$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}. \tag{62}$$

10. Conjugate functions in the plane

For the transformation of Laplace's operator in two dimensions we have similarly

$$\begin{aligned}\partial_x^2+\partial_y^2 &= A\partial_u^2+2H\partial_u\partial_v+B\partial_v^2+(u_{xx}+u_{yy})\partial_u+(v_{xx}+v_{yy})\partial_v \\ &= J\left\{\partial_u\left(\frac{A}{J}\partial_u+\frac{H}{J}\partial_v\right)+\partial_v\left(\frac{H}{J}\partial_u+\frac{B}{J}\partial_v\right)\right\},\end{aligned} \tag{63}$$

where $$A = u_x^2+u_y^2, \qquad H = u_x v_x+u_y v_y, \qquad B = v_x^2+v_y^2,$$
$$J = u_x v_y-u_y v_x.$$

Let us now determine, if possible, the coordinates (u, v) in which the corresponding Laplace's *equation*

$$\frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2} = 0 \tag{64}$$

retains its form: that is to say, in which the transformation (63) is

$$\partial_x^2+\partial_y^2 = \lambda(\partial_u^2+\partial_v^2)$$

for some λ. The conditions necessary and sufficient for this are

$$A = B, \qquad H = 0 \tag{65}$$

$$u_{xx}+u_{yy} = 0, \qquad v_{xx}+v_{yy} = 0. \tag{66}$$

The last *three* conditions tell us that u, v are necessarily orthogonal solutions of Laplace's equation (64). To consider the full effect of the *four* conditions let us, however, begin with the two relations (65), i.e.

$$u_x^2+u_y^2 = v_x^2+v_y^2, \qquad u_x v_x+u_y v_y = 0,$$

which give

$$\frac{u_x}{v_y} = -\frac{v_x}{u_y} = \pm\sqrt{\frac{u_x^2-v_x^2}{v_y^2-u_y^2}} = \pm 1.$$

The choice of sign in ± 1 means merely a choice between u and v. Take the positive sign and so write

$$u_x = v_y, \qquad u_y = -v_x. \tag{67}$$

If between these two equations we eliminate u and v in turn, we at once get the two equations (66). Thus the four conditions (65), (66) for the invariance of Laplace's equation are redundant and reduce to two independent conditions: but not to less than two, since the two equations (66), for instance, are evidently independent. The two equations (67) represent the most convenient form of these two independent conditions and we therefore enunciate:

(68) *Laplace's two-dimensional equation*

$$\frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2} = 0$$

is invariant for a change of coordinates from (x, y) *to* (u, v), *where* $u(x,y)$, $v(x,y)$ *satisfy the pair of differential equations*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x};$$

and only for a change to such coordinates.

These equations (67) are known as the *Cauchy-Riemann equations* and are fundamental in the theory of Functions of a Complex Variable. A pair of functions u, v that satisfy the Cauchy-Riemann equations are called *conjugate functions*. We might therefore restate (68) by saying that Laplace's equation is invariant for transformation to coordinates that are conjugate functions of x, y and to such coordinates alone.

An interpretation of the Cauchy-Riemann equations can be given in terms of complex number: that interpretation here must necessarily be

purely formal, since our theory has been throughout a theory of real functions of a real variable. Write, then,

$$w \equiv u+iv,$$

where i is the unit of imaginary number. We may combine the two Cauchy-Riemann equations in the form

$$u_x+iv_x+i(u_y+iv_y) = 0,$$

which now may be written formally as

$$\frac{\partial w}{\partial x}+i\frac{\partial w}{\partial y} = 0.$$

Again, a formal solution of this differential equation is

$$w = f(x+iy).$$

Thus we may formally define a pair of conjugate functions by the equation

$$u+iv = f(x+iy). \tag{69}$$

This equation, of course, carries with it the conjugate equation

$$u-iv = f(x-iy),$$

if we suppose f to be constructed out of purely real elements, so that we may write explicitly

$$u = \frac{f(x+iy)+f(x-iy)}{2}, \qquad v = \frac{f(x+iy)-f(x-iy)}{2i}.$$

The further implications of this theory and, in particular, the meaning to be ascribed to $f(x+iy)$ a 'function' of a complex variable are fundamental to the theory of Functions of a Complex Variable. We shall find, however, when f is an elementary function, e.g. a polynomial or exponential, that the implications of the formal relation (69) can be substantiated in real terms without much difficulty.

11. Invariance of Laplace's equation in space

We go on to inquire whether conjugate functions can be similarly defined in three dimensions. The transformation of Laplace's operator is now, it will be remembered,

$$\nabla^2 = A\partial_u^2+B\partial_v^2+C\partial_w^2+2F\partial_v\partial_w+2G\partial_w\partial_u+2H\partial_u\partial_v+$$
$$+(\nabla^2 u)\partial_u+(\nabla^2 v)\partial_v+(\nabla^2 w)\partial_w.$$

The conditions necessary and sufficient for invariance are now

$$A = B = C, \tag{70}$$

$$F = 0, \qquad G = 0, \qquad H = 0, \tag{71}$$

$$\nabla^2 u = 0, \qquad \nabla^2 v = 0, \qquad \nabla^2 w = 0. \tag{72}$$

Since, by (71), the coordinates are orthogonal, we may write (70) as

$$U^2 = V^2 = W^2 = 1/\theta^2, \quad \text{say}. \tag{73}$$

By (60) we may write (72) as

$$\frac{\partial}{\partial u}\left(\frac{U}{VW}\right) = 0, \qquad \frac{\partial}{\partial v}\left(\frac{V}{WU}\right) = 0, \qquad \frac{\partial}{\partial w}\left(\frac{W}{UV}\right) = 0,$$

i.e.

$$\frac{\partial \theta}{\partial u} = 0, \qquad \frac{\partial \theta}{\partial v} = 0, \qquad \frac{\partial \theta}{\partial w} = 0,$$

by (73). Thus θ is a mere constant. Going over now to u, v, w as fundamental coordinates we have

$$x_v x_w + y_v y_w + z_v z_w = 0, \quad \text{etc.}, \tag{74}$$

for the orthogonal conditions, and

$$x_u^2 + y_u^2 + z_u^2 = \text{constant}, \quad \text{etc.}, \tag{75}$$

from (73). Differentiation of (74) in u gives

$$\sum x_{uv} x_w + \sum x_{wu} x_v = 0,$$

where the summation $\sum$ is over x, y, z. Similar differentiation of the two other orthogonal conditions gives

$$\sum x_{vw} x_u + \sum x_{uv} x_w = 0,$$
$$\sum x_{wu} x_v + \sum x_{vw} x_u = 0,$$

whence $\quad \sum x_{vw} x_u = 0, \qquad \sum x_{wu} x_v = 0, \qquad \sum x_{uv} x_w = 0.$

Differentiation of (75) in v and in w gives

$$\sum x_{uv} x_u = 0, \qquad \sum x_{uw} x_u = 0$$

with four similar equations. The three equations

$$\sum x_{vw} x_u = 0, \qquad \sum x_{vw} x_v = 0, \qquad \sum x_{vw} x_w = 0$$

give

$$x_{vw} = 0, \qquad y_{vw} = 0, \qquad z_{vw} = 0, \tag{76}$$

since

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} \neq 0.$$

Similarly, $\quad x_{wu},\ y_{wu},\ z_{wu}, \qquad x_{uv},\ y_{uv},\ z_{uv}$

are all zero.

Again, differentiation in u of (75) and the two orthogonal conditions $\sum x_u x_v = 0$, $\sum x_u x_w = 0$ gives

$$\sum x_{uu} x_u = 0,$$
$$\sum x_{uu} x_v = 0, \quad \text{since } x_{uv} = 0, \text{ etc.},$$
$$\sum x_{uu} x_w = 0, \quad \text{since } x_{uw} = 0, \text{ etc.}$$

Hence also we must have

$$x_{uu} = 0, \qquad y_{uu} = 0, \qquad z_{uu} = 0,$$

since $$\frac{\partial(x,y,z)}{\partial(u,v,w)} \neq 0.$$

Similarly, x_{vv}, y_{vv}, z_{vv}, x_{ww}, y_{ww}, z_{ww} all vanish. Thus every second derivative of x, y, z in u, v, w vanishes, and the transformation is a *linear* one. We may therefore write the transformation in the form

$$\left.\begin{aligned} x &= l_1 u + m_1 v + n_1 w \\ y &= l_2 u + m_2 v + n_2 w \\ z &= l_3 u + m_3 v + n_3 w \end{aligned}\right\}, \tag{77}$$

where $l_1, m_1, \ldots$ are all constants. The conditions (74), (75) are now

$$m_1 n_1 + m_2 n_2 + m_3 n_3 = 0, \quad \text{etc.},$$
$$l_1^2 + l_2^2 + l_3^2 = m_1^2 + m_2^2 + m_3^2 = n_1^2 + n_2^2 + n_3^2 = k^2, \quad \text{say.}$$

Thus (x, y, z) transform to (ku, kv, kw) by the standard transformation of rectangular cartesian coordinates, and so (u, v, w) themselves are also rectangular cartesian coordinates, but possibly with a changed unit of length. We may therefore say that

(78) *Laplace's equation in three dimensions is invariant only for a change to other rectangular axes.*

The foregoing arguments extend, it can be seen, to the generalization of Laplace's equation to any dimensions

$$\sum \frac{\partial^2 V}{\partial x_r^2} = 0.$$

The existence of conjugate coordinates other than rectangular cartesians thus appears as a peculiarity of two dimensions. The analytical clue to this peculiarity is to be found in the behaviour of the terms A/J, B/J, which in two dimensions become U/V, V/U, i.e. unity, under the conditions $U^2 = V^2$. Thus the pair of conditions for invariance

$$\frac{\partial}{\partial u}\left(\frac{A}{J}\right) = 0, \qquad \frac{\partial}{\partial v}\left(\frac{B}{J}\right) = 0$$

are satisfied automatically. This reduction in the number of conditions leaves room for the existence of curvilinear conjugate functions in two dimensions.

12. Hamilton's dual transformation

Given $f(x, y, z)$ a differentiable function of (x, y, z) we may consider the special transformation to coordinates

$$u \equiv \frac{\partial f}{\partial x}, \qquad v \equiv \frac{\partial f}{\partial y}, \qquad w \equiv \frac{\partial f}{\partial z}. \tag{79}$$

It is prerequisite that the Jacobian of the transformation do not vanish: this Jacobian is the Hessian

$$\begin{vmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial x \partial z} \\ \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial y^2} & \dfrac{\partial^2 f}{\partial y \partial z} \\ \dfrac{\partial^2 f}{\partial x \partial z} & \dfrac{\partial^2 f}{\partial y \partial z} & \dfrac{\partial^2 f}{\partial z^2} \end{vmatrix}, \tag{80}$$

an expression we have already considered in the theory of maxima and minima. We shall shortly turn to discuss it as the Jacobian of this transformation. For the moment we merely suppose it not to vanish.

We began, in the field (x, y, z), with a special dependent variable f. Let us consider also, in the transformed field, the special dependent variable

$$F(u, v, w) \equiv xu+yv+zw-f(x, y, z). \tag{81}$$

Total differentiation gives

$$DF = x\,Du + y\,Dv + z\,Dw + u\,Dx + v\,Dy + w\,Dz - Df.$$

But, by definition of u, v, w,

$$Df = u\,Dx + v\,Dy + w\,Dz.$$

Hence
$$DF = x\,Du + y\,Dv + z\,Dw, \tag{82}$$

and so, since F is expressly to be considered a function of u, v, w, we have

$$x = \frac{\partial F}{\partial u}, \qquad y = \frac{\partial F}{\partial v}, \qquad z = \frac{\partial F}{\partial w}. \tag{83}$$

The complete transformation may therefore be given as

$$\left.\begin{array}{c} u = \dfrac{\partial f}{\partial x}, \qquad v = \dfrac{\partial f}{\partial y}, \qquad w = \dfrac{\partial f}{\partial z}, \\ F+f = xu+yu+zw, \\ x = \dfrac{\partial F}{\partial u}, \qquad y = \dfrac{\partial F}{\partial v}, \qquad z = \dfrac{\partial F}{\partial w} \end{array}\right\}. \tag{84}$$

It is evidently symmetrical as between the two sets of variables (x, y, z), (u, v, w), and is conveniently called the *dual* or *reciprocal* transformation: it is also known as *Hamilton's transformation*. The transformations in two and three independent variables can be shown to correspond to geometrical reciprocation: in the plane with regard to a certain parabola and in space with regard to a certain paraboloid.

In extension of (83) we may go on to express the second derivatives

of F in terms of those of f. Total differentiation of, say, the first equation of (83) gives

$$Dx = \frac{\partial^2 F}{\partial u^2} Du + \frac{\partial^2 F}{\partial u \partial v} Dv + \frac{\partial^2 F}{\partial u \partial w} Dw. \tag{85}$$

Total differentiation of (79) gives similarly, in briefer notation,

$$\begin{aligned} Du &= f_{xx} Dx + f_{xy} Dy + f_{xz} Dz, \\ Dv &= f_{xy} Dx + f_{yy} Dy + f_{yz} Dz, \\ Dw &= f_{xz} Dx + f_{yz} Dy + f_{zz} Dz. \end{aligned}$$

Solving these equations for Dx we have

$$H\,Dx = \begin{vmatrix} f_{yy} & f_{yz} \\ f_{yz} & f_{zz} \end{vmatrix} Du + \begin{vmatrix} f_{yz} & f_{zz} \\ f_{xy} & f_{xz} \end{vmatrix} Dv + \begin{vmatrix} f_{xy} & f_{xz} \\ f_{yy} & f_{yz} \end{vmatrix} Dw, \tag{86}$$

where H is the Hessian (80). The coefficients of the differentials on the right are, of course, minors in H and may be written as Jacobians of a pair of f_x, f_y, f_z, just as H itself is the Jacobian of all three. By a principle already invoked we may identify (85), (86) and so get

$$\left.\begin{aligned} H\frac{\partial^2 F}{\partial u^2} &= \frac{\partial(f_y, f_z)}{\partial(y,z)} \\ H\frac{\partial^2 F}{\partial u \partial v} &= \frac{\partial(f_y, f_z)}{\partial(z,x)} = \frac{\partial(f_z, f_x)}{\partial(y,z)} \end{aligned}\right\}. \tag{87}$$

13. Vanishing of the Hessian

The dual transformation becomes improper, if the Hessian vanishes, for then the u, v, w of (79) are no longer independent. Bearing in mind the analogy of the vanishing of the Jacobian, we inquire the meaning of a vanishing Hessian. Alternatively, we may regard $H = 0$ as a differential equation defining f and seek its solution.

If the Hessian vanishes, then u, v, w are functionally connected, as we have just said, and we may write

$$w = \psi(u, v). \tag{88}$$

We may still define an F by (81), though it is no longer a function of three independent variables, and we still deduce from that definition the differential relation (82),

$$\begin{aligned} DF &= x\,Du + y\,Dv + z\,Dw \\ &= (x + z\psi_u)\,Du + (y + z\psi_v)\,Dv, \quad \text{by (88)}, \\ &= \lambda\,Du + \mu\,Dv, \quad \text{say, for brevity.} \end{aligned}$$

This gives at once

$$F_x = \lambda u_x + \mu v_x, \qquad F_y = \lambda u_y + \mu v_y, \qquad F_z = \lambda u_z + \mu v_z,$$

and so
$$\frac{\partial(F,u,v)}{\partial(x,y,z)} = 0.$$

Thus F also is a function of u, v only, say $F = \chi(u,v)$, and we have from (81)
$$f = xu+yv+z\psi(u,v)-\chi(u,v), \tag{89}$$
where
$$\frac{\partial f}{\partial x} = u, \qquad \frac{\partial f}{\partial y} = v, \qquad \frac{\partial f}{\partial z} = \psi(u,v),$$
and ψ, χ are two arbitrary functions (equal in number to the order of the differential equation, as we should expect). It is, however, better to increase the apparent number of arbitrary functions† and write more symmetrically
$$f = x\theta(u,v)+y\phi(u,v)+z\psi(u,v)-\chi(u,v), \tag{90}$$
where now $\quad f_x = \theta(u,v), \qquad f_y = \phi(u,v), \qquad f_z = \psi(u,v).$

But actual differentiation gives
$$f_x = \theta(u,v)+u_x\frac{\partial f}{\partial u}+v_x\frac{\partial f}{\partial v}, \quad \text{etc.}$$
We must therefore have
$$u_x\frac{\partial f}{\partial u}+v_x\frac{\partial f}{\partial v} = 0, \qquad u_y\frac{\partial f}{\partial u}+v_y\frac{\partial f}{\partial v} = 0, \qquad u_z\frac{\partial f}{\partial u}+v_z\frac{\partial f}{\partial v} = 0. \tag{91}$$
These give

either (i)
$$\frac{\partial f}{\partial u} = 0, \qquad \frac{\partial f}{\partial v} = 0,$$

or (ii)
$$\left\|\begin{matrix} u_x & u_y & u_z \\ v_x & v_y & v_z \end{matrix}\right\| = 0.$$

Under (ii), x, y, z can all be eliminated between u, v, say $v = \omega(u)$. Then θ, ϕ, ψ, χ all become functions of a single variable u, and we may most simply imagine that v is merely suppressed in (90). In this case equations (91) become
$$u_x\frac{\partial f}{\partial u} = 0, \qquad u_y\frac{\partial f}{\partial u} = 0, \qquad u_z\frac{\partial f}{\partial u} = 0.$$
These again give

either (ii a)
$$\frac{\partial f}{\partial u} = 0,$$

or (ii b) $\qquad u_x = 0, \qquad u_y = 0, \qquad u_z = 0.$

Under (ii b), u is a mere constant and it is simplest to suppose that both u, v are suppressed in (90).

† We do not, of course, increase their actual number.

Thus all three cases (i), (ii *a*), (ii *b*) are covered by the general form (90) with the added equations

$$\frac{\partial f}{\partial u} = 0, \qquad \frac{\partial f}{\partial v} = 0, \tag{92}$$

which may be regarded as equations to determine u, v when suitable forms have been given to the arbitrary functions θ, ϕ, ψ, χ. We get (ii *a*) by choosing arbitrary functions independent of v: the second equation of (92) is then satisfied automatically, but the first remains to determine the single unknown u. We get (ii *b*) by choosing functions independent of u, v, i.e. mere constants: both equations (92) are then satisfied automatically and there remain no unknowns u, v to be determined.

The asymmetric form (89) first obtained for f does not lend itself so well to such a comprehensive statement. We sum up our results in the enunciation:

(93) *If the Hessian of $f(x,y,z)$ vanishes identically, then f must be of the form*

$$f = x\theta(u,v)+y\phi(u,v)+z\psi(u,v)-\chi(u,v),$$

where u, v are defined by the equations

$$x\theta_u+y\phi_u+z\psi_u-\chi_u = 0, \qquad x\theta_v+y\phi_v+z\psi_v-\chi_v = 0.$$

14. The bordered Hessian

If, in (90), we define a new set of arbitrary functions

$$\lambda \equiv \theta/\chi, \qquad \mu \equiv \phi/\chi, \qquad \nu \equiv \psi/\chi, \qquad \rho \equiv -1/\chi,$$

we can write the relation between f, x, y, z in the form, symmetrical in x, y, z, f,

$$1 = x\lambda(u,v)+y\mu(u,v)+z\nu(u,v)+f\rho(u,v),$$

where

$$0 = x\lambda_u+y\mu_u+z\nu_u+f\rho_u = x\lambda_v+y\mu_v+z\nu_v+f\rho_v.$$

The relation given by the vanishing of the Hessian of f with respect to x, y, z is therefore symmetrical in x, y, z, f.

To prove this directly we discuss the 'bordered' Hessian, and to abbreviate the working we now diminish by one the number of independent variables. Consider, then, the function of two independent variables $z = z(x,y)$ defined implicitly by the relation $f(x,y,z) = 0$. We shall prove the following theorem connecting the Hessian of $z(x,y)$ with the bordered Hessian of $f(x,y,z)$:

(94) *If the function* $z = z(x,y)$ *is defined implicitly by the relation* $f(x,y,z) = 0$, *then*

$$\begin{vmatrix} z_{xx} & z_{xy} \\ z_{xy} & z_{yy} \end{vmatrix} = -f_z^{-4} \begin{vmatrix} f_{xx} & f_{xy} & f_{xz} & f_x \\ f_{xy} & f_{yy} & f_{yz} & f_y \\ f_{xz} & f_{yz} & f_{zz} & f_z \\ f_x & f_y & f_z & 0 \end{vmatrix}.$$

This theorem may be deduced as a particular case of the more general theorem (20) of chapter X. It is, however, useful for present purposes to give an independent proof.

Denote the Hessian and the bordered Hessian above by

$$H(z;x,y) \quad \text{and} \quad H'(f;x,y,z) \quad \text{respectively.}$$

We then have by manipulation of rows

$$H'(f;x,y,z) = f_z^2 \begin{vmatrix} \partial_x(f_x/f_z) & \partial_y(f_x/f_z) & \partial_z(f_x/f_z) & 0 \\ \partial_x(f_y/f_z) & \partial_y(f_y/f_z) & \partial_z(f_y/f_z) & 0 \\ f_{xz} & f_{yz} & f_{zz} & f_z \\ f_x & f_y & f_z & 0 \end{vmatrix},$$

i.e.

$$H'(f;x,y,z) = -f_z^3 \begin{vmatrix} \partial_x(f_x/f_z) & \partial_y(f_x/f_z) & \partial_z(f_x/f_z) \\ \partial_x(f_y/f_z) & \partial_y(f_y/f_z) & \partial_z(f_y/f_z) \\ f_x & f_y & f_z \end{vmatrix}, \tag{95}$$

$$= -f_z^3 \begin{vmatrix} (\partial_x+z_x\partial_z)(f_x/f_z) & (\partial_y+z_y\partial_z)(f_x/f_z) & \partial_z(f_x/f_z) \\ (\partial_x+z_x\partial_z)(f_y/f_z) & (\partial_y+z_y\partial_z)(f_y/f_z) & \partial_z(f_y/f_z) \\ 0 & 0 & f_z \end{vmatrix},$$

by manipulation of columns.

Now in the field of the two independent variables x, y we have

$$\frac{\partial}{\partial x} = \partial_x+z_x\partial_z, \qquad \frac{\partial}{\partial y} = \partial_y+z_y\partial_z,$$

$$\frac{\partial z}{\partial x} = -f_x/f_z, \qquad \frac{\partial z}{\partial y} = -f_y/f_z.$$

Thus $$H'(f;x,y,z) = -f_z^4 \begin{vmatrix} \dfrac{\partial^2 z}{\partial x^2} & \dfrac{\partial^2 z}{\partial x \partial y} \\ \dfrac{\partial^2 z}{\partial x \partial y} & \dfrac{\partial^2 z}{\partial y^2} \end{vmatrix} = -f_z^4 H(z;x,y),$$

which establishes the theorem.

We see, then, that the relation

$$H(z;x,y) = 0$$

implies the relation $\qquad H'(f;x,y,z) = 0,$

which is symmetrical in x, y, z. This confirms the point made at the beginning of the section.

More generally, with n variables, the theorem takes the form

(96) *If $x_1,\ldots,x_n$ are connected by the relation $f(x_1,\ldots,x_n) = 0$, then*

$$H(x_1;x_2,\ldots,x_n) = (-)^n H'(f;x_1,\ldots,x_n)/f_1^{n+1}.$$

We consider, finally, the interpretation of the condition

$$H'(f;x,y,z) = 0,$$

regarded as a differential equation in f. By (95), $H'(f;x,y,z) = 0$ gives

$$\frac{\partial(f_x/f_z, f_y/f_z, f)}{\partial(x,y,z)} = 0. \tag{97}$$

Again, write $\qquad F \equiv xf_x+yf_y+zf_z.$

Then

$$F_x = xf_{xx}+yf_{xy}+zf_{xz}+f_x,$$
$$F_y = xf_{xy}+yf_{yy}+zf_{yz}+f_y,$$
$$F_z = xf_{xz}+yf_{yz}+zf_{zz}+f_z.$$

Thus, by manipulation of columns,

$$H'(f;x,y,z) = \begin{vmatrix} f_{xx} & f_{xy} & f_{xz} & F_x \\ f_{xy} & f_{yy} & f_{yz} & F_y \\ f_{xz} & f_{yz} & f_{zz} & F_z \\ f_x & f_y & f_z & F \end{vmatrix}.$$

Manipulation of rows, as in the proof of (94), then gives

$$H'(f;x,y,z) = -f_z^4 F\frac{\partial(f_x/f_z, f_y/f_z, F/f_z)}{\partial(x,y,z)},$$

so that $H'(f;x,y,z) = 0$ leads to

$$\frac{\partial(f_x/f_z, f_y/f_z, F/f_z)}{\partial(x,y,z)} = 0. \tag{98}$$

From (97), (98) we see that

$$f, \qquad f_x/f_z, \qquad f_y/f_z, \qquad F/f_z$$

are all functions of the same two parameters. Take f itself as one of these and let u denote the other. We can therefore write

$$\frac{f_x}{\theta(f,u)} = \frac{f_y}{\phi(f,u)} = \frac{f_z}{\psi(f,u)} = \frac{F}{\chi(f,u)}, \tag{99}$$

where θ, ϕ, ψ, χ are arbitrary functions. Substitution for F gives

$$x\theta(f,u)+y\phi(f,u)+z\psi(f,u) = \chi(f,u).$$

Differentiation in x, in y, and in z gives

$$\theta+\lambda f_x+\mu u_x = 0,$$
$$\phi+\lambda f_y+\mu u_y = 0,$$
$$\psi+\lambda f_z+\mu u_z = 0,$$

where $\quad \lambda = x\theta_f+y\phi_f+z\psi_f-\chi_f, \qquad \mu = x\theta_u+y\phi_u+z\psi_u-\chi_u.$

Substitution for θ, ϕ, ψ from (99) gives

$$\mu \begin{Vmatrix} f_x & f_y & f_z \\ u_x & u_y & u_z \end{Vmatrix} = 0,$$

i.e. either $\mu = 0$ or else x, y, z can be eliminated between f and u, i.e. u is a function of f, and θ, ϕ, ψ, χ are functions of f only. This is merely a particular case of the general form and gives $\mu = 0$ automatically. We may therefore sum up our result into the theorem:

(100) *If the bordered Hessian*

$$\begin{vmatrix} f_{xx} & f_{xy} & f_{xz} & f_x \\ f_{xy} & f_{yy} & f_{yz} & f_y \\ f_{xz} & f_{yz} & f_{zz} & f_z \\ f_x & f_y & f_z & 0 \end{vmatrix}$$

vanishes identically, then $f(x,y,z)$ is of a form given implicitly by

$$x\theta(f,u)+y\phi(f,u)+z\psi(f,u) = \chi(f,u),$$

where θ, ϕ, ψ, χ are arbitrary functions and u is defined in terms of x, y, z, f by the relation $\quad x\theta_u(f,u)+y\phi_u(f,u)+z\psi_u(f,u) = \chi_u(f,u).$

WORKED EXAMPLE

Transform the operator $\qquad \dfrac{\partial^3}{\partial x^3}+\dfrac{\partial^3}{\partial y^3}$

to curvilinear coordinates (u, v).

As in § 9 above, we may write

$$\begin{aligned} \Delta &\equiv \frac{\partial^3}{\partial x^3}+\frac{\partial^3}{\partial y^3} \\ &= \alpha_0\frac{\partial^3}{\partial u^3}+3\alpha_1\frac{\partial^3}{\partial u^2\partial v}+3\alpha_2\frac{\partial^3}{\partial u\partial v^2}+\alpha_3\frac{\partial^3}{\partial v^3}+ \\ &\qquad +3\Big(\beta_0\frac{\partial^2}{\partial u^2}+2\beta_1\frac{\partial^2}{\partial u\partial v}+\beta_2\frac{\partial^2}{\partial v^2}\Big)+\gamma_0\frac{\partial}{\partial u}+\gamma_1\frac{\partial}{\partial v}, \end{aligned}$$

where $\quad \alpha_0 = \sum u_x^3, \quad \alpha_1 = \sum u_x^2 v_x, \quad \alpha_2 = \sum u_x v_x^2, \quad \alpha_3 = \sum v_x^3, \qquad (1)$

($\sum$ denoting summation over x, y), and β_0, β_1, β_2, γ_0, γ_1 have still to be determined.

Operation with Δ on u, v, u^2, uv, v^2 in turn gives

$$\Delta u = \gamma_0, \qquad \Delta v = \gamma_1,$$
$$\Delta u^2 = 6\beta_0 + 2u\gamma_0, \qquad \Delta v^2 = 6\beta_2 + 2v\gamma_1,$$
$$\Delta uv = 6\beta_1 + v\gamma_0 + u\gamma_1.$$

But
$$\Delta u^2 = 2u\Delta u + 6\sum u_{xx}u_x,$$
$$\Delta uv = u\Delta v + v\Delta u + 3\sum (u_{xx}v_x + u_x v_{xx}),$$
$$\Delta v^2 = 2v\Delta v + 6\sum v_{xx}v_x.$$

Thus
$$\left.\begin{aligned} \beta_0 &= \sum u_{xx}u_x \\ 2\beta_1 &= \sum (u_{xx}v_x + u_x v_{xx}) \\ \beta_2 &= \sum v_{xx}v_x \end{aligned}\right\}. \tag{2}$$

Suppose now that λ, μ, X, Y are any functions of x, y such that
$$\lambda = Xu_x + Yu_y, \qquad \mu = Xv_x + Yv_y. \tag{3}$$

Then
$$X = \left(\frac{\lambda}{J}\right)v_y - \left(\frac{\mu}{J}\right)u_y, \qquad Y = -\left(\frac{\lambda}{J}\right)v_x + \left(\frac{\mu}{J}\right)u_x,$$

where $J \equiv \dfrac{\partial(u,v)}{\partial(x,y)}$. Thus
$$\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} = \left(v_y\frac{\partial}{\partial x} - v_x\frac{\partial}{\partial y}\right)\left(\frac{\lambda}{J}\right) + \left(-u_y\frac{\partial}{\partial x} + u_x\frac{\partial}{\partial y}\right)\left(\frac{\mu}{J}\right),$$

the other terms cancelling.

That is to say,
$$\frac{1}{J}\left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y}\right) = \frac{\partial}{\partial u}\left(\frac{\lambda}{J}\right) + \frac{\partial}{\partial v}\left(\frac{\mu}{J}\right). \tag{4}$$

In (3) put in succession

(i) $X = u_x^2$, $Y = u_y^2$, then $\lambda = \alpha_0$, $\mu = \alpha_1$;

(ii) $X = u_x v_x$, $Y = u_y v_y$, then $\lambda = \alpha_1$, $\mu = \alpha_2$;

(iii) $X = v_x^2$, $Y = v_y^2$, then $\lambda = \alpha_2$, $\mu = \alpha_3$.

We have correspondingly from (4), using (2),
$$\left.\begin{aligned} \frac{2\beta_0}{J} &= \frac{2}{J}\sum u_{xx}u_x = \frac{\partial}{\partial u}\left(\frac{\alpha_0}{J}\right) + \frac{\partial}{\partial v}\left(\frac{\alpha_1}{J}\right) \\ \frac{2\beta_1}{J} &= \frac{2}{J}\sum (v_{xx}u_x + u_{xx}v_x) = \frac{\partial}{\partial u}\left(\frac{\alpha_1}{J}\right) + \frac{\partial}{\partial v}\left(\frac{\alpha_2}{J}\right) \\ \frac{2\beta_2}{J} &= \frac{2}{J}\sum v_{xx}v_x = \frac{\partial}{\partial u}\left(\frac{\alpha_2}{J}\right) + \frac{\partial}{\partial v}\left(\frac{\alpha_3}{J}\right) \end{aligned}\right\}, \tag{5}$$

Write now
$$2\omega = \sum (v_{xx}u_x - u_{xx}v_x), \tag{6}$$

then
$$\beta_0 = \sum u_{xx}u_x, \qquad \beta_1 - \omega = \sum u_{xx}v_x,$$
and
$$\beta_1 + \omega = \sum v_{xx}u_x, \qquad \beta_2 = \sum v_{xx}v_x.$$

Thus, again by (3), (4),
$$\left.\begin{aligned} \frac{\Delta u}{J} &= \frac{\partial}{\partial u}\left(\frac{\beta_0}{J}\right) + \frac{\partial}{\partial v}\left(\frac{\beta_1 - \omega}{J}\right) \\ \frac{\Delta v}{J} &= \frac{\partial}{\partial u}\left(\frac{\beta_1 + \omega}{J}\right) + \frac{\partial}{\partial v}\left(\frac{\beta_2}{J}\right) \end{aligned}\right\}. \tag{7}$$

So far, then,

$$\frac{1}{J}\Delta = \frac{1}{J}(\alpha_0, \alpha_1, \alpha_2, \alpha_3)\left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v}\right)^3 + \frac{2}{J}(\beta_0, \beta_1, \beta_2)\left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v}\right)^2 +$$
$$+\frac{\partial}{\partial u}\left(\frac{\beta_0}{J}\frac{\partial}{\partial u}+\frac{\beta_1}{J}\frac{\partial}{\partial v}\right)+\frac{\partial}{\partial v}\left(\frac{\beta_1}{J}\frac{\partial}{\partial u}+\frac{\beta_2}{J}\frac{\partial}{\partial v}\right)+\frac{\partial}{\partial u}\left(\frac{\omega}{J}\right)\frac{\partial}{\partial v}-\frac{\partial}{\partial v}\left(\frac{\omega}{J}\right)\frac{\partial}{\partial u}, \text{ by (7)},$$
$$= \frac{\partial}{\partial u}\left\{\frac{\alpha_0}{J}\frac{\partial^2}{\partial u^2}+\frac{2\alpha_1}{J}\frac{\partial^2}{\partial u\partial v}+\frac{\alpha_2}{J}\frac{\partial^2}{\partial v^2}-\frac{\beta_0}{J}\frac{\partial}{\partial u}+\frac{\beta_1}{J}\frac{\partial}{\partial v}\right\}+$$
$$+\frac{\partial}{\partial v}\left\{\frac{\alpha_1}{J}\frac{\partial^2}{\partial u^2}+\frac{2\alpha_2}{J}\frac{\partial^2}{\partial u\partial v}+\frac{\alpha_3}{J}\frac{\partial^2}{\partial v^2}+\frac{\beta_1}{J}\frac{\partial}{\partial u}+\frac{\beta_2}{J}\frac{\partial}{\partial v}\right\}+\frac{\partial}{\partial u}\left(\frac{\omega}{J}\right)\frac{\partial}{\partial v}-\frac{\partial}{\partial v}\left(\frac{\omega}{J}\right)\frac{\partial}{\partial u}, \text{ by (5)}.$$

Hence, again by (5), we may write

$$\frac{2}{J}\Delta = \frac{\partial}{\partial u}\left(\frac{\alpha_0}{J}\frac{\partial^2}{\partial u^2}+\frac{2\alpha_1}{J}\frac{\partial^2}{\partial u\partial v}+\frac{\alpha_2}{J}\frac{\partial^2}{\partial v^2}\right)+\frac{\partial}{\partial v}\left(\frac{\alpha_1}{J}\frac{\partial^2}{\partial u^2}+\frac{2\alpha_2}{J}\frac{\partial^2}{\partial u\partial v}+\frac{\alpha_3}{J}\frac{\partial^2}{\partial v^2}\right)-$$
$$-\frac{\partial^2}{\partial u^2}\left(\frac{\alpha_0}{J}\frac{\partial}{\partial u}+\frac{\alpha_1}{J}\frac{\partial}{\partial v}\right)+\frac{\partial^2}{\partial u\partial v}\left(\frac{\alpha_1}{J}\frac{\partial}{\partial u}+\frac{\alpha_2}{J}\frac{\partial}{\partial v}\right)-\frac{\partial^2}{\partial v^2}\left(\frac{\alpha_2}{J}\frac{\partial}{\partial u}+\frac{\alpha_3}{J}\frac{\partial}{\partial v}\right)+$$
$$+\frac{\partial}{\partial u}\left(\frac{2\omega}{J}\right)\frac{\partial}{\partial v}-\frac{\partial}{\partial v}\left(\frac{2\omega}{J}\right)\frac{\partial}{\partial u} \qquad (8)$$

It remains to consider ω. No very simple expression for ω in terms of $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ seems to exist. We shall be content with one or two equivalent forms which recall the coefficients in the covariants H and G of the binary cubic

$$(\alpha_0, \alpha_1, \alpha_2, \alpha_3)(X, Y)^3.$$

We find, on differentiating and cancelling, that

$$u_x^2\frac{\partial}{\partial x}(v_x v_y)-u_x v_x\frac{\partial}{\partial x}(u_x v_y+u_y v_x)+v_x^2\frac{\partial}{\partial x}(u_x u_y) = (u_x v_{xx}-u_{xx}v_x)J,$$

and similarly, on interchanging x and y, that

$$u_y^2\frac{\partial}{\partial y}(v_x v_y)-u_y v_y\frac{\partial}{\partial y}(u_x v_y+u_y v_x)+v_y^2\frac{\partial}{\partial y}(u_x u_y) = -(u_y v_{yy}-u_{yy}v_y)J.$$

Thus

$$2\omega J = \left(u_x^2\frac{\partial}{\partial x}-u_y^2\frac{\partial}{\partial y}\right)v_x v_y-\left(u_x v_x\frac{\partial}{\partial x}-u_y v_y\frac{\partial}{\partial y}\right)(u_x v_y+u_y v_x)+$$
$$+\left(v_x^2\frac{\partial}{\partial x}-v_y^2\frac{\partial}{\partial y}\right)u_x u_y \qquad (9)$$

Write

$$\left.\begin{aligned} A_0 &= \alpha_0\alpha_2-\alpha_1^2 = -u_x u_y J^2 \\ 2A_1 &= \alpha_0\alpha_3-\alpha_1\alpha_2 = (u_x v_y+u_y v_x)J^2 \\ A_2 &= \alpha_1\alpha_3-\alpha_2^2 = -v_x v_y J^2 \end{aligned}\right\} \qquad (10)$$

Then

$$2\omega J = \left(u_x^2\frac{\partial}{\partial x}-u_y^2\frac{\partial}{\partial y}\right)\left(\frac{A_2}{J^2}\right)-2\left(u_x v_x\frac{\partial}{\partial x}-u_y v_y\frac{\partial}{\partial y}\right)\left(\frac{A_1}{J^2}\right)+\left(v_x^2\frac{\partial}{\partial x}-v_y^2\frac{\partial}{\partial y}\right)\left(\frac{A_0}{J^2}\right).$$

Again, from (1),

$$\alpha_0 v_y-\alpha_1 u_y = u_x^2 J, \qquad \alpha_0 v_x-\alpha_1 u_x = -u_y^2 J,$$
$$\alpha_1 v_y-\alpha_2 u_y = u_x v_x J, \qquad \alpha_1 v_x-\alpha_2 u_x = -u_y v_y J,$$
$$\alpha_2 v_y-\alpha_3 u_y = v_x^2 J, \qquad \alpha_2 v_x-\alpha_3 u_x = -v_y^2 J.$$

Hence, writing

$$\partial = J^2\left(u_y\frac{\partial}{\partial x}+u_x\frac{\partial}{\partial y}\right), \qquad \partial' = J^2\left(v_y\frac{\partial}{\partial x}+v_x\frac{\partial}{\partial y}\right),$$

we have

$$2\omega J^4 = (\alpha_0\partial'-\alpha_1\partial)\Big(\frac{A_2}{J^2}\Big)-2(\alpha_1\partial'-\alpha_2\partial)\Big(\frac{A_1}{J^2}\Big)+(\alpha_2\partial'-\alpha_3\partial)\Big(\frac{A_0}{J^2}\Big).$$

Again, from (10) we get

$$-v_x A_0+u_x A_1 = \tfrac{1}{2}u_x J^3, \qquad -v_x A_1+u_x A_2 = \tfrac{1}{2}v_x J^3,$$
$$v_y A_0-u_y A_1 = \tfrac{1}{2}u_y J^3, \qquad v_y A_1-u_y A_2 = \tfrac{1}{2}v_y J^3,$$

and
$$A_1^2-A_0A_2 = \tfrac{1}{4}J^6.$$

Thus
$$\tfrac{1}{2}J\partial = A_0\Big(v_y\frac{\partial}{\partial x}-v_x\frac{\partial}{\partial y}\Big)+A_1\Big(-u_y\frac{\partial}{\partial x}+u_x\frac{\partial}{\partial y}\Big)$$
$$-J\Big(A_0\frac{\partial}{\partial u}+A_1\frac{\partial}{\partial v}\Big),$$
$$\tfrac{1}{2}J\partial' = A_1\Big(v_y\frac{\partial}{\partial x}-v_x\frac{\partial}{\partial y}\Big)+A_2\Big(-u_y\frac{\partial}{\partial x}+u_x\frac{\partial}{\partial y}\Big)$$
$$-J\Big(A_1\frac{\partial}{\partial u}+A_2\frac{\partial}{\partial v}\Big).$$

Hence, at length,

$$2\omega J^4 = (\alpha_0\partial'-\alpha_1\partial)\Big(\frac{A_2}{J^2}\Big)-2(\alpha_1\partial'-\alpha_2\partial)\Big(\frac{A_1}{J^2}\Big)+(\alpha_2\partial'-\alpha_3\partial)\Big(\frac{A_0}{J^2}\Big), \tag{11}$$

where
$$\partial = -2\Big(A_0\frac{\partial}{\partial u}+A_1\frac{\partial}{\partial v}\Big), \qquad \partial' = -2\Big(A_1\frac{\partial}{\partial u}+A_2\frac{\partial}{\partial v}\Big).$$

Substituting for ∂, ∂' we may also write this as

$$2\omega J^4 = \Big(C_0\frac{\partial}{\partial u}+C_1\frac{\partial}{\partial v}\Big)\Big(\frac{A_2}{J^2}\Big)-2\Big(C_1\frac{\partial}{\partial u}+C_2\frac{\partial}{\partial v}\Big)\Big(\frac{A_1}{J^2}\Big)+\Big(C_2\frac{\partial}{\partial u}+C_3\frac{\partial}{\partial v}\Big)\Big(\frac{A_0}{J^2}\Big), \tag{12}$$

where
$$C_0 = \alpha_0^2\alpha_3-3\alpha_0\alpha_1\alpha_2-2\alpha_1^3, \qquad C_1 = \alpha_0\alpha_1\alpha_3-2\alpha_0\alpha_2^2+\alpha_1^2\alpha_2,$$
$$C_2 = \alpha_0\alpha_2\alpha_3-2\alpha_1^2\alpha_3+\alpha_1\alpha_2^2, \qquad C_3 = \alpha_0\alpha_3^2-3\alpha_1\alpha_2\alpha_3+2\alpha_2^3.$$

Since, by (10),
$$\left.\begin{aligned}\alpha_0A_2-2\alpha_1A_1+\alpha_2A_0 &= 0\\ \alpha_1A_2-2\alpha_2A_1+\alpha_3A_0 &= 0\end{aligned}\right\}, \tag{13}$$

we may rewrite (11) in the form

$$2\omega J^5 = -A_2\partial'\Big(\frac{\alpha_0}{J}\Big)+(A_2\partial+2A_1\partial')\Big(\frac{\alpha_1}{J}\Big)-(2A_1\partial+A_0\partial')\Big(\frac{\alpha_2}{J}\Big)+A_0\partial\Big(\frac{\alpha_3}{J}\Big),$$

i.e., at full length,

$$\begin{aligned}4(A_1^2-A_0A_2)\omega = {} & -A_1A_2\frac{\partial}{\partial u}\Big(\frac{\alpha_0}{J}\Big)-A_2^2\frac{\partial}{\partial v}\Big(\frac{\alpha_0}{J}\Big)+\\ & +(2A_1^2+A_0A_2)\frac{\partial}{\partial u}\Big(\frac{\alpha_1}{J}\Big)+3A_1A_2\frac{\partial}{\partial v}\Big(\frac{\alpha_1}{J}\Big)-\\ & -3A_0A_1\frac{\partial}{\partial u}\Big(\frac{\alpha_2}{J}\Big)-(2A_1^2+A_0A_2)\frac{\partial}{\partial v}\Big(\frac{\alpha_2}{J}\Big)+\\ & +A_0^2\frac{\partial}{\partial u}\Big(\frac{\alpha_3}{J}\Big)+A_0A_1\frac{\partial}{\partial v}\Big(\frac{\alpha_3}{J}\Big)\end{aligned} \tag{14}$$

Now (13) shows that the *operator*

$$A_2(\alpha_0\partial'-\alpha_1\partial)-2A_1(\alpha_1\partial'-\alpha_2\partial)+A_0(\alpha_2\partial'-\alpha_3\partial)$$

vanishes identically. That is to say, the operator

$$A_2\left(C_0\frac{\partial}{\partial u}+C_1\frac{\partial}{\partial v}\right)-2A_1\left(C_1\frac{\partial}{\partial u}-C_2\frac{\partial}{\partial v}\right)-A_0\left(C_2\frac{\partial}{\partial u}+C_3\frac{\partial}{\partial v}\right)$$

vanishes identically. We have therefore the identical relations

$$A_2C_0-2A_1C_1-A_0C_2=0$$

$$A_2C_1+2A_1C_2-A_0C_3=0,$$

and we may rewrite (12) in the form

$$2\omega J^3=A_2\frac{\partial}{\partial u}\left(\frac{C_0}{J^3}\right)+\left(2A_1\frac{\partial}{\partial u}-A_2\frac{\partial}{\partial v}\right)\left(\frac{C_1}{J^3}\right)+$$
$$+\left(A_0\frac{\partial}{\partial u}-2A_1\frac{\partial}{\partial v}\right)\left(\frac{C_2}{J^3}\right)+A_0\frac{\partial}{\partial v}\left(\frac{C_3}{J^3}\right). \tag{15}$$

By using in (8) one or other of the foregoing forms for ω we at length complete the transformation of Δ into curvilinear coordinates (u,v).

In (12), (14), (15) we can, of course, simplify the formulae for ω by transferring J and its powers from the denominators on the right, past the operators, and across to the left: this is permissible in virtue of the identical vanishing of the operator noted above. It is in better conformity, however, with (8) and with the analogous transformation of ∇^2 to state results in terms α_0/J, α_1/J,... rather than in the α's themselves.

By imposing a suitable linear transformation on the variables x, y, we can transform the cubic operator

$$\frac{\partial^3}{\partial x^3}+\frac{\partial^3}{\partial y^3}$$

into the general cubic operator with constant coefficients

$$a_0\frac{\partial^3}{\partial x^3}+3a_1\frac{\partial^3}{\partial x^2\partial y}+3a_2\frac{\partial^3}{\partial x\partial y^2}+a_3\frac{\partial^3}{\partial y^3}. \tag{16}$$

The expressions for α_0, α_1, α_2, α_3 are correspondingly changed. We can define them most completely by the identity of cubics

$$(\alpha_0,\alpha_1,\alpha_2,\alpha_3)(1,t)^3=(a_0,a_1,a_2,a_3)(u_x+tv_x,u_y+tv_y)^3,$$

where t is arbitrary.

The definitions of A_0, A_1, A_2, namely,

$$A_0=\alpha_0\alpha_2-\alpha_1^2,\qquad 2A_1=\alpha_0\alpha_3-\alpha_1\alpha_2,\qquad A_2=\alpha_1\alpha_3-\alpha_2^2,$$

remain unchanged; so too do the definitions of C_0, C_1, C_2, C_3 in terms of α_0, α_1, α_2, α_3. With these changes the formulae (8), (11), (12), (14), (15) become applicable to the more general cubic operator (16).

EXAMPLES XII

1. For transformation between coordinates (u, v) and (x, y), show that the relations between differential operators

$$\left.\begin{aligned} u\frac{\partial}{\partial u}-v\frac{\partial}{\partial v} &= x\frac{\partial}{\partial x}-y\frac{\partial}{\partial y} \\ v\frac{\partial}{\partial u}+u\frac{\partial}{\partial v} &= y\frac{\partial}{\partial x}+x\frac{\partial}{\partial y} \end{aligned}\right\}$$

lead to the relations between differentials

$$\left.\begin{aligned} x\,du+y\,dv &= u\,dx+v\,dy \\ y\,du-x\,dv &= v\,dx-u\,dy \end{aligned}\right\}.$$

Show also that the most general transformation of coordinates compatible with these relations is

$$u = ax, \qquad v = ay,$$

where a is some constant.

2. For transformation between coordinates (u, v) and (x, y), prove the equivalence of the relations

$$\left.\begin{aligned} u\frac{\partial}{\partial u}+v\frac{\partial}{\partial v} &= x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y} \\ v\frac{\partial}{\partial u}-u\frac{\partial}{\partial v} &= y\frac{\partial}{\partial x}-x\frac{\partial}{\partial y} \end{aligned}\right\}$$

to the relations

$$\left.\begin{aligned} y\,du+x\,dv &= v\,dx+u\,dy \\ x\,du-y\,dv &= u\,dx-v\,dy \end{aligned}\right\},$$

and show that the only transformation of coordinates compatible with these relations is

$$u = ax-by, \qquad v = bx+ay,$$

where a, b are constants.

3. Show that the equivalence of operators

$$\left.\begin{aligned} uv\left(u\frac{\partial}{\partial u}+v\frac{\partial}{\partial v}\right) &= xy\left(x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}\right) \\ v\frac{\partial}{\partial u}-u\frac{\partial}{\partial v} &= y\frac{\partial}{\partial x}-x\frac{\partial}{\partial y} \end{aligned}\right\}$$

gives the equivalence of differentials

$$\left.\begin{aligned} \frac{v\,du-u\,dv}{u^2+v^2} &= \frac{y\,dx-x\,dy}{x^2+y^2} \\ \frac{u\,du+v\,dv}{uv(u^2+v^2)} &= \frac{x\,dx+y\,dy}{xy(x^2+y^2)} \end{aligned}\right\},$$

and that $u(x, y)$, $v(x, y)$ are then defined by

$$\frac{u}{x} = \frac{v}{y} = \frac{a}{\sqrt{(a^2-x^2-y^2)}},$$

where a is a constant.

4. Show that the equivalences of operators

$$\left.\begin{aligned}
&\text{(i)} && u\frac{\partial}{\partial u}+v\frac{\partial}{\partial v}=x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y} \\
& && uv\left(v\frac{\partial}{\partial u}-u\frac{\partial}{\partial v}\right)=xy\left(y\frac{\partial}{\partial x}-x\frac{\partial}{\partial y}\right)
\end{aligned}\right\},$$

$$\left.\begin{aligned}
&\text{(ii)} && (u+v)\left(\frac{\partial}{\partial u}+\frac{\partial}{\partial v}\right)=(x+y)\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}\right) \\
& && \frac{\partial}{\partial u}-\frac{\partial}{\partial v}=\frac{\partial}{\partial x}-\frac{\partial}{\partial y}
\end{aligned}\right\},$$

$$\left.\begin{aligned}
&\text{(iii)} && (u-v)\left(\frac{\partial}{\partial u}-\frac{\partial}{\partial v}\right)=(x-y)\left(\frac{\partial}{\partial x}-\frac{\partial}{\partial y}\right) \\
& && (u+v)\left(\frac{\partial}{\partial u}+\frac{\partial}{\partial v}\right)=(x+y)\left(\frac{\partial}{\partial x}-\frac{\partial}{\partial y}\right)
\end{aligned}\right\}$$

lead respectively to the transformation-formulae

$$\text{(i)}\quad \frac{u}{x^a}=\frac{v}{by^a}=\sqrt{\frac{a(x^2+y^2)}{x^{2a}+b^2y^{2a}}},$$

$$\text{(ii)}\quad \left.\begin{aligned} u&=(a+1)x+ay+b \\ v&=ax+(a+1)y-b \end{aligned}\right\},$$

$$\text{(iii)}\quad \left.\begin{aligned} u&=ax+by \\ v&=bx+ay \end{aligned}\right\},$$

where a, b are arbitrary constants.

5. For transformation between coordinates (u,v) and (x,y), prove that

$$\frac{\partial u}{\partial x}\frac{\partial x}{\partial u}=\frac{\partial v}{\partial y}\frac{\partial y}{\partial v}.$$

For transformation between coordinates $(u_1,\ldots,u_n)$ and $(x_1,\ldots,x_n)$, prove that

$$\text{(i)}\quad \frac{\partial(u_1,\ldots,u_r)}{\partial(x_1,\ldots,x_r)}=\frac{\partial(u_1,\ldots,u_n)}{\partial(x_1,\ldots,x_n)}\frac{\partial(x_{r+1},\ldots,x_n)}{\partial(u_{r+1},\ldots,u_n)},$$

$$\text{(ii)}\quad \frac{\partial(u_1,\ldots,u_r)}{\partial(x_1,\ldots,x_r)}\frac{\partial(x_1,\ldots,x_r)}{\partial(u_1,\ldots,u_r)}=\frac{\partial(u_{r+1},\ldots,u_n)}{\partial(x_{r+1},\ldots,x_n)}\frac{\partial(x_{r+1},\ldots,x_n)}{\partial(u_{r+1},\ldots,u_n)},$$

$$\text{(iii)}\quad \frac{\partial(u_1,\ldots,u_r)}{\partial(x_1,\ldots,x_r)}\frac{\partial(x_1,\ldots,x_s)}{\partial(u_1,\ldots,u_s)}=\frac{\partial(u_{s+1},\ldots,u_n)}{\partial(x_{s+1},\ldots,x_n)}\frac{\partial(x_{r+1},\ldots,x_n)}{\partial(u_{r+1},\ldots,u_n)}.$$

6. Establish the following propositions by arguments in the field of the real variable:

(i) if u, v are conjugate functions of x, y, and u', v' are also conjugate functions of x, y, then $u+u'$, $v+v'$ are conjugate functions of x, y and so also are $uu'-vv'$, $uv'-u'v$;

(ii) if u, v are conjugate functions of x, y, then x, y are conjugate functions of u, v;

(iii) if u, v are conjugate functions of x, y and if U, V are conjugate functions of u, v, then U, V are conjugate functions of x, y.

7. Show, by arguments in the field of the real variable, that the functions

$$u = \frac{1+x}{1+2x+x^2+y^2}, \qquad u = \frac{\cos x \cosh y}{\cos 2x + \cosh 2y}, \qquad u = \tan^{-1}\frac{2x}{1-x^2-y^2}$$

are 'harmonic' functions, i.e. that they satisfy the two-dimensional Laplace's equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

If u is a harmonic function, show that

$$\tan^{-1}\left(\frac{\partial u}{\partial x}\Big/\frac{\partial u}{\partial y}\right)$$

is also a harmonic function.

Find the functions v conjugate to the above harmonic functions u.

8. (i) If $u(x,y), f(x,y)$ satisfy Laplace's equation in two dimensions, show that

$$u_x, \qquad xu_x + yu_y, \qquad yu_x - xu_y, \qquad x^2u_{xx} + 2xyu_{xy} + y^2u_{yy},$$

and $$f(u_x, u_y)$$

also satisfy the equation.

(ii) If $u(x,y,z)$ satisfies Laplace's equation in three dimensions, show that

$$u_x, \qquad xu_x + yu_y + zu_z, \qquad yu_x - xu_y, \qquad y^2u_{xx} - 2xyu_{xy} + x^2u_{yy} + zu_z,$$

and the quadratic $(u_{xx}, u_{yy}, u_{zz}, u_{yz}, u_{zx}, u_{xy})(x, y, z)^2$

also satisfy the equation.

If $F(u_x, u_y, u_z)$ satisfies the equation whenever u does, show that F is merely a linear function of its arguments.

9. (i) For the transformation of coordinates

$$u = \frac{x}{x^2+y^2+z^2}, \qquad v = \frac{y}{x^2+y^2+z^2}, \qquad w = \frac{z}{x^2+y^2+z^2},$$

show that Laplace's equation $\nabla^2 V = 0$ becomes

$$(u^2+v^2+w^2)\left(\frac{\partial^2 V}{\partial u^2} + \frac{\partial^2 V}{\partial v^2} + \frac{\partial^2 V}{\partial w^2}\right) = 2\left(u\frac{\partial V}{\partial u} + v\frac{\partial V}{\partial v} + w\frac{\partial V}{\partial w}\right).$$

(ii) Transform Laplace's equation to 'confocal' coordinates u, v, w, namely the roots of the cubic in θ,

$$\frac{x^2}{a^2+\theta} + \frac{y^2}{b^2+\theta} + \frac{z^2}{c^2+\theta} = 1.$$

10. (i) If $$\frac{\partial^2 z}{\partial x \partial y}\Big/\frac{\partial z}{\partial x}\frac{\partial z}{\partial y}$$

is a function of z alone, show that

$$u = \frac{\partial z}{\partial x}\Big/\frac{\partial z}{\partial y}$$

satisfies the differential equation

$$u\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial u}{\partial x}\frac{\partial u}{\partial y}.$$

(ii) The variables x, y, z are connected by an implicit relation $f(x,y,z) = 0$ which is such that

$$\frac{\partial^2 z}{\partial x \partial y}\Big/\frac{\partial z}{\partial x}\frac{\partial z}{\partial y}$$

is a function of z only Show that

$$\frac{\partial^2 x}{\partial y \partial z} \Big/ \frac{\partial x}{\partial y}\frac{\partial x}{\partial z}$$

is a function of x only, and that f itself satisfies the differential equation

$$\frac{\partial(f, f_{xx}/f_x)}{\partial(y,z)} + \frac{\partial(f, f_{yy}/f_y)}{\partial(z,x)} + \frac{\partial(f, f_{zz}/f_z)}{\partial(x,y)} = 0$$

11 If
$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) f(u) = 0,$$

show that $\theta = \log(u_x/u_y)$ satisfies the differential equation

$$\frac{\partial^2\theta}{\partial x^2} + \frac{\partial^2\theta}{\partial y^2} = \tanh\theta\left\{\left(\frac{\partial\theta}{\partial x}\right)^2 + \left(\frac{\partial\theta}{\partial y}\right)^2\right\}$$

and find the differential equation satisfied by u itself

If $u(x, y)$, $v(x, y)$ form a system of orthogonal coordinates such that $\nabla^2 f(u) = 0$, show that we can find $g(v)$ such that $\nabla^2 g(v) = 0$ and that $f(u)$, $g(v)$ are conjugate functions

12 If u is defined as a function of $x_1, \ldots, x_n$ by the equation

$$\text{(i)} \qquad \frac{x_1^2}{a_1 - u} + \cdots + \frac{x_n^2}{a_n - u} = 1,$$

where $a_1, \ldots, a_n$ are constants, show that some function $V = V(u)$ can be found to satisfy the generalized Laplace's equation

$$\text{(ii)} \qquad \frac{\partial^2 V}{\partial x_1^2} + \cdots + \frac{\partial^2 V}{\partial x_n^2} = 0,$$

and determine this function V as an integral or otherwise

If $u_1, \ldots, u_n$ are the n roots of (i) regarded as an equation in u, and $V_1, \ldots, V_n$ are the corresponding values of $V(u)$ show that the differential equation (ii) may be written as

$$\frac{1}{(u_1 - u_2) \cdots (u_1 - u_n)}\frac{\partial^2 V}{\partial V_1^2} + \cdots + \frac{1}{(u_n - u_1) \cdots (u_n - u_{n-1})}\frac{\partial^2 V}{\partial V_n^2} = 0$$

13 If
$$u = \frac{x_1^2}{x_1^2 + \cdots + x_n^2},$$

show that, for a suitable value of p there is a function $V(u)$ that satisfies the generalized Laplace's equation

$$\frac{\partial^2 V}{\partial x_1^2} + \cdots + \frac{\partial^2 V}{\partial x_n^2} = 0$$

and determine the appropriate p and $V(u)$

14 (i) If u is defined as a function of x, y by the equation

$$\frac{x^2}{a^2 + u} + \frac{y^2}{b^2 + u} = 1,$$

determine a function $f(u)$ such that

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) f(u) = 0$$

(ii) If u is defined as a function of x, y by the equation

$$x^2 \operatorname{sech}^2 u - y^2 \operatorname{cosech}^2 u = 1,$$

show that it satisfies the differential equation

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial y^2},$$

and express u in the form $\qquad f(x+y)+g(x-y)$.

15. Show that functions of

(i) $$1+x^2-xy+y^2+x^2y^2+(x-y)(1-xy),$$

(ii) $$\log(1-2xy)-\log(1-x-y)$$

can be found to satisfy the differential equation

$$\frac{\partial^2 z}{\partial x \partial y} = 0,$$

and express (i), (ii) in the form $F\{\phi(x)+\psi(y)\}$.

16. Show that functions of

$$\text{(i)}\ \frac{x^2+y^2+1}{2x}, \qquad \text{(ii)}\ \frac{x^2-y^2+1}{2xy},$$

$$\text{(iii)}\ \frac{\cos^2 x+\cosh^2 y}{2\cos x\cosh y}, \qquad \text{(iv)}\ \frac{\sin^2 x+\cosh^2 y}{2\sin x\cosh y}, \qquad \text{(v)}\ \frac{\sin x}{\cosh y},$$

can be found to satisfy Laplace's equation in two dimensions. Determine the corresponding conjugate functions.

17. If u is defined as a function of $x_1,\ldots,x_n$ by the equation

$$\frac{x_1^2}{u+a_1}+\ldots+\frac{x_n^2}{u+a_n} = 1,$$

show that a function $V = V(u)$ can be found to satisfy the differential equation

$$\frac{\partial^2 V}{\partial x_1^2}+\ldots+\frac{\partial^2 V}{\partial x_n^2} = \frac{b_1}{x_1}\frac{\partial V}{\partial x_1}+\ldots+\frac{b_n}{x_n}\frac{\partial V}{\partial x_n},$$

where $a_1,\ldots,a_n, b_1,\ldots,b_n$ are constants. Obtain this function in explicit terms when for every r, $\qquad$ (i) $b_r = -1$; (ii) $b_r = 1$; (iii) $b_r = 3$.

18. (i) If u is defined as a function of x, y, z by the equation

$$\frac{x^2}{u}+\frac{y^2}{u+a}+\frac{z^2}{u-a} = 1,$$

show that $V = \sec^{-1}(u/a)$ satisfies the differential equation

$$\frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2}+\frac{\partial^2 V}{\partial z^2}+\frac{1}{x}\frac{\partial V}{\partial x} = 0,$$

and $V = \cosh^{-1}(u/a)$ the differential equation

$$\frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2}+\frac{\partial^2 V}{\partial z^2}-\frac{1}{x}\frac{\partial V}{\partial x} = 0.$$

(ii) If u is defined as a function of x_1, x_2, x_3, x_4, x_5 by the equation

$$\frac{x_1^2}{u}+\frac{x_2^2}{u+a}+\frac{x_3^2}{u-a}+\frac{x_4^2}{u+b}+\frac{x_5^2}{u-b} = 1,$$

show that $V = \log\{\sqrt{(u^2-a^2)}+\sqrt{(u^2-b^2)}\}$ satisfies the differential equation

$$\frac{\partial^2 V}{\partial x_1^2}+\frac{\partial^2 V}{\partial x_2^2}+\frac{\partial^2 V}{\partial x_3^2}+\frac{\partial^2 V}{\partial x_4^2}+\frac{\partial^2 V}{\partial x_5^2} = \frac{3}{x_1}\frac{\partial V}{\partial x_1},$$

and $$V = \log\frac{a\sqrt{(u^2-b^2)}-b\sqrt{(u^2-a^2)}}{a\sqrt{(u^2-b^2)}+b\sqrt{(u^2-a^2)}}$$

satisfies the differential equation

$$\frac{\partial^2 V}{\partial x_1^2}+\frac{\partial^2 V}{\partial x_2^2}+\frac{\partial^2 V}{\partial x_3^2}+\frac{\partial^2 V}{\partial x_4^2}+\frac{\partial^2 V}{\partial x_5^2}+\frac{1}{x_1}\frac{\partial V}{\partial x_1} = 0.$$

19. If $$\nabla^2 = \frac{\partial^2}{\partial x_1^2}+\ldots+\frac{\partial^2}{\partial x_n^2}, \qquad r^2 \equiv x_1^2+\ldots+x_n^2,$$

and if $f(x_1,\ldots,x_n)$ is homogeneous of degree p in its arguments and also a solution of the equation $\nabla^2 f = 0$, show that, for a given positive integer m, we can find $s(m)$ such that $$(\nabla^2)^{m-1} r^s f \neq 0, \qquad (\nabla^2)^m r^s f = 0.$$

20. (i) If $$\nabla^2 = \frac{\partial^2}{\partial x_1^2}+\ldots+\frac{\partial^2}{\partial x_n^2}, \qquad r^2 = x_1^2+\ldots+x_n^2,$$

and $f(x_1,\ldots,x_n)$ is a solution of the equation

$$\nabla^2 f = 0,$$

show that a solution of the equation

$$\nabla^2 F = r^p f(x_1,\ldots,x_n)$$

is $$F = \frac{r^{p+2}}{2p+4}\int\limits_a^1 f(x_1 t,\ldots,x_n t)t^{\frac{1}{2}p+\frac{1}{2}n-1}\,dt,$$

provided that $t = a$ is a zero of

$$f(x_1 t,\ldots,x_n t)t^{\frac{1}{2}p+\frac{1}{2}n};$$

and show also that $\nabla^2(r^{-p-2}F) = 0$.

(ii) Show that any solution of

$$(\nabla^2)^m U = 0$$

can be written as $$U = u_1+r^2u_2+r^4u_3+\ldots+r^{2m-2}u_m,$$

where $u_1, u_2,\ldots, u_m$ are solutions of $\nabla^2 u = 0$.

21. Show that

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)^m (x^2+y^2)^n = \begin{cases} 0 & (m > n), \\ \left[\dfrac{n!\,2^m}{(n-m)!}\right]^2 (x^2+y^2)^{n-m} & (m \leq n), \end{cases}$$

and evaluate similarly

(i) $$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)^m (x^2+y^2+z^2)^n,$$

(ii) $$\left(\frac{\partial^3}{\partial x^3}+\frac{\partial^3}{\partial y^3}+\frac{\partial^3}{\partial z^3}-3\frac{\partial^3}{\partial x\partial y\partial z}\right)^m (x^3+y^3+z^3-3xyz)^n,$$

(iii) $$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}-\frac{\partial^2}{\partial y\partial z}-\frac{\partial^2}{\partial z\partial x}-\frac{\partial^2}{\partial x\partial y}\right)^m (x^2+y^2+z^2-yz-zx-xy)^n,$$

(iv) $$\begin{vmatrix} \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \cdot & \dfrac{\partial}{\partial x_p} \\ \dfrac{\partial}{\partial x_p} & \dfrac{\partial}{\partial x_1} & \cdot & \dfrac{\partial}{\partial x_{p-1}} \\ \cdot & \cdot & \cdot & \cdot \\ \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3} & \cdot & \dfrac{\partial}{\partial x_1} \end{vmatrix}^m \begin{vmatrix} x_1 & x_2 & \cdot & x_p \\ x_p & x_1 & \cdot & x_{p-1} \\ \cdot & \cdot & \cdot & \cdot \\ x_2 & x_3 & \cdot & x_1 \end{vmatrix}^n.$$

22. Show that

$$\left(\frac{\partial}{\partial x}\right)^n F\left(\frac{x^2+y^2+z^2}{2}\right)$$
$$= \left\{x^n D^n + \frac{n(n-1)}{2} x^{n-2} D^{n-1} + \frac{n(n-1)(n-2)(n-3)}{2.4} x^{n-4} D^{n-2} + \ldots\right\} F\left(\frac{x^2+y^2+z^2}{2}\right),$$

where D denotes differentiation of F with respect to its argument.

If $\psi(x,y,z)$ denotes a homogeneous polynomial of degree n, show that

$$\psi\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) F\left(\frac{x^2+y^2+z^2}{2}\right)$$
$$= \left\{D^n + \frac{\nabla^2 D^{n-1}}{2} + \frac{(\nabla^2)^2 D^{n-2}}{2.4} + \frac{(\nabla^2)^3 D^{n-3}}{2.4.6} + \ldots\right\} \psi(x,y,z) F\left(\frac{x^2+y^2+z^2}{2}\right),$$

with the convention that ∇^2 operates only on ψ, and, as above, D only on F.

23. (i) If the families of surfaces

$$u(x,y,z) = \text{constant}, \qquad v(x,y,z) = \text{constant}$$

are both cut orthogonally by a third family of surfaces $w(x,y,z) =$ constant, show that the functions u, v are connected by the differential relation

$$\begin{vmatrix} U(v_x)-V(u_x) & U(v_y)-V(u_y) & U(v_z)-V(u_z) \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix} = 0,$$

where U, V denote the operators

$$U = u_x\frac{\partial}{\partial x} + u_y\frac{\partial}{\partial y} + u_z\frac{\partial}{\partial z}, \qquad V = v_x\frac{\partial}{\partial x} + v_y\frac{\partial}{\partial y} + v_z\frac{\partial}{\partial z}.$$

(ii) If the surfaces $u, v, w =$ constant are mutually orthogonal, show that u (and therefore also v and w) satisfies the differential equation

$$\begin{vmatrix} U(u_{xx})-\frac{1}{3}\theta_{xx} & u_{xx} & 1 & 2u_x & 0 & 0 \\ U(u_{yy})-\frac{1}{3}\theta_{yy} & u_{yy} & 1 & 0 & 2u_y & 0 \\ U(u_{zz})-\frac{1}{3}\theta_{zz} & u_{zz} & 1 & 0 & 0 & 2u_z \\ U(u_{yz})-\frac{1}{3}\theta_{yz} & u_{yz} & 0 & 0 & u_z & u_y \\ U(u_{zx})-\frac{1}{3}\theta_{zx} & u_{zx} & 0 & u_z & 0 & u_x \\ U(u_{xy})-\frac{1}{3}\theta_{xy} & u_{xy} & 0 & u_y & u_x & 0 \end{vmatrix} = 0,$$

where the operator U is defined above and $\theta \equiv u_x^2+u_y^2+u_z^2$.

Show that we may replace the above determinant by a determinant of the third order in which a typical column is

$$u_z^2\{U(u_{yy})-\tfrac{1}{3}\theta_{yy}\} - 2u_y u_z\{U(u_{yz})-\tfrac{1}{3}\theta_{yz}\} + u_y^2\{U(u_{zz})-\tfrac{1}{3}\theta_{zz}\}$$
$$u_z^2 u_{yy} - 2u_y u_z u_{yz} + u_y^2 u_{zz}$$
$$u_y^2 + u_z^2.$$

24. If u, v, w are a set of orthogonal parameters in space of three dimensions, show that

$$\frac{\partial x}{\partial u}\frac{\partial^2 x}{\partial v\partial w} + \frac{\partial y}{\partial u}\frac{\partial^2 y}{\partial v\partial w} + \frac{\partial z}{\partial u}\frac{\partial^2 z}{\partial v\partial w} = 0.$$

Show that the differential equation of lines of curvature on a surface $u =$ constant, namely

$$\begin{vmatrix} dx & dy & dz \\ d\left(\dfrac{\partial u}{\partial x}\right) & d\left(\dfrac{\partial u}{\partial y}\right) & d\left(\dfrac{\partial u}{\partial z}\right) \\ \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} & \dfrac{\partial u}{\partial z} \end{vmatrix} = 0,$$

can be written $\qquad dv\,dw = 0.\qquad$ [Dupin's theorem.]

25. The 'measure of curvature' K on a surface $u(x,y,z) =$ constant is given by

$$K = -\begin{vmatrix} u_{xx} & u_{xy} & u_{xz} & u_x \\ u_{xy} & u_{yy} & u_{yz} & u_y \\ u_{xz} & u_{yz} & u_{zz} & u_z \\ u_x & u_y & u_z & 0 \end{vmatrix} \div (u_x^2+u_y^2+u_z^2)^2.$$

If (v, w) are parameters (not necessarily orthogonal) on the surface, show that

$$K = (BC-F^2) \div (bc-f^2)^2, \quad \text{where}$$

$$B = \begin{vmatrix} x_{vv} & y_{vv} & z_{vv} \\ x_v & y_v & z_v \\ x_w & y_w & z_w \end{vmatrix}, \quad C = \begin{vmatrix} x_{ww} & y_{ww} & z_{ww} \\ x_v & y_v & z_v \\ x_w & y_w & z_w \end{vmatrix}, \quad F = \begin{vmatrix} x_{vw} & y_{vw} & z_{vw} \\ x_v & y_v & z_v \\ x_w & y_w & z_w \end{vmatrix},$$

and $\quad b = x_v^2+y_v^2+z_v^2, \quad c = x_w^2+y_w^2+z_w^2, \quad f = x_v x_w+y_v y_w+z_v z_w.$

Show further that

$$4fH^2K = \frac{\partial H}{\partial v}\frac{\partial H}{\partial w} + \left(f\frac{\partial c}{\partial v} - c\frac{\partial b}{\partial w}\right)\frac{\partial H}{\partial v} + \left(f\frac{\partial b}{\partial w} - b\frac{\partial c}{\partial v}\right)\frac{\partial H}{\partial w} -$$

$$+ H\left\{\frac{\partial b}{\partial w}\frac{\partial c}{\partial v} - \frac{\partial b}{\partial v}\frac{\partial c}{\partial w} - 2f\left(\frac{\partial^2 b}{\partial w^2} + \frac{\partial^2 c}{\partial v^2} - 2\frac{\partial^2 f}{\partial v\,\partial w}\right)\right\},$$

where $H = bc-f^2$.

26. In the 'dual' transformation (84) of § 12, if x, y, z and therefore u, v, w be expressed in terms of a third set of coordinates (ξ, η, ζ), show that

$$\frac{\partial(x,u)}{\partial(\xi,\eta)} + \frac{\partial(y,v)}{\partial(\xi,\eta)} + \frac{\partial(z,w)}{\partial(\xi,\eta)} = 0.$$

27. If p, q, r, s, t have their usual meanings of partial derivatives of z with respect to x, y, and if new variables X, Y, Z are defined by the relations

$$aX+hY+g = p, \qquad hX+bY+f = q,$$

$$Z+z = (ax+hy+g)X+(hx+by+f)Y+(gx+fy+c),$$

where a, b, c, f, g, h are constants such that $ab \neq h^2$, show that we have

$$P = ax+hy+g, \qquad Q = hx+by+f,$$

$$\frac{bR-2hS+aT}{br-2hs+at} = \frac{ab-h^2}{rt-s^2} = \frac{RT-S^2}{ab-h^2},$$

$$a(St-sT)-h(Rt-Tr)+b(Rs-Sr) = 0,$$

$$\frac{hR-aS}{hr-as} = \frac{bS-hT}{bs-ht} = \frac{aT-bR}{at-br},$$

where P, Q, R, S, T denote the partial derivatives of Z with respect to X, Y.

28. If $u(x,y,z)$ is homogeneous of the first degree, show that its Hessian vanishes.

If
$$X \equiv \frac{\partial u}{\partial x}, \qquad Y \equiv \frac{\partial u}{\partial y}, \qquad Z \equiv \frac{\partial u}{\partial z},$$
show that
$$\frac{\partial Z}{\partial X} = -\frac{x}{z}, \qquad \frac{\partial Z}{\partial Y} = -\frac{y}{z},$$
and
$$\frac{\partial^2 Z}{\partial X^2} = -\lambda \frac{\partial^2 u}{\partial y^2}, \qquad \frac{\partial^2 Z}{\partial X \partial Y} = \lambda \frac{\partial^2 u}{\partial x \partial y}, \qquad \frac{\partial^2 Z}{\partial Y^2} = -\lambda \frac{\partial^2 u}{\partial x^2},$$
where
$$\lambda^{-1} = z\left\{\frac{\partial^2 u}{\partial x^2}\frac{\partial^2 u}{\partial y^2} - \left(\frac{\partial^2 u}{\partial x \partial y}\right)^2\right\}$$

If, further, u satisfies the partial differential equation
$$au_{xx} + bu_{yy} + cu_{zz} + 2fu_{yz} + 2gu_{zx} + 2hu_{xy} = 0,$$
obtain the corresponding partial differential equation satisfied by Z.

29. (i) If $u(x,y,z)$ is homogeneous of degree n ($n \neq 1$) and its Hessian vanishes, show that it can be written in the form
$$u = \{xf(t) + y\,g(t) + z\,h(t)\}^n, \qquad 0 = xf'(t) + y\,g'(t) + z\,h'(t).$$
Express in this form a homogeneous quadratic that is the product of two linear factors.

(ii) If $u(x,y,z)$ is homogeneous of degree n ($n \neq 1$) and its Hessian does not vanish, and if
$$X = \frac{\partial u}{\partial x}, \qquad Y = \frac{\partial u}{\partial y}, \qquad Z = \frac{\partial u}{\partial z},$$
show that
$$x = (n-1)\frac{\partial u}{\partial X}, \qquad y = (n-1)\frac{\partial u}{\partial Y}, \qquad z = (n-1)\frac{\partial u}{\partial Z}$$
and
$$H\frac{\partial^2 u}{\partial X^2} = \begin{vmatrix} \dfrac{\partial^2 u}{\partial y^2} & \dfrac{\partial^2 u}{\partial y \partial z} \\[2ex] \dfrac{\partial^2 u}{\partial y \partial z} & \dfrac{\partial^2 u}{\partial z^2} \end{vmatrix}, \qquad H\frac{\partial^2 u}{\partial X \partial Y} = -\begin{vmatrix} \dfrac{\partial^2 u}{\partial x \partial y} & \dfrac{\partial^2 u}{\partial x \partial z} \\[2ex] \dfrac{\partial^2 u}{\partial y \partial z} & \dfrac{\partial^2 u}{\partial z^2} \end{vmatrix}$$

30. If $u(x,y,z)$ is homogeneous in x, y, z, and if $v(x,y,z)$ is a function of u only, show that $H(v)/H(u)$ is a function of u only.

31. If $u(x,y,z)$ satisfies the partial differential equation
$$F(u_x, u_y, u_z) = 0,$$
where F is a polynomial with constant coefficients, show that the Hessian of u vanishes.

If the coordinates of a point on the surface $F(X,Y,Z) = 0$ can be written parametrically in the form
$$X = f(t,t'), \qquad Y = g(t,t'), \qquad Z = h(t,t'),$$
show that the general solution of the above differential equation can be written in the form
$$u = xf(t,t') + y\,g(t,t') + z\,h(t,t') + \phi(t,t'),$$
$$\frac{\partial u}{\partial t} = 0, \qquad \frac{\partial u}{\partial t'} = 0,$$
where ϕ is an arbitrary function.

32. If
$$\begin{vmatrix} u_{xx} & u_{xy} & u_x \\ u_{xy} & u_{yy} & u_y \\ u_x & u_y & u \end{vmatrix} = 0,$$
show that u is of the form
$$u = \exp\{x\,\phi(t)+y\,\psi(t)+\chi(t)\},$$
where
$$0 = x\phi'(t)+y\,\psi'(t)+\chi'(t).$$

More generally, if
$$\begin{vmatrix} u_{xx} & u_{xy} & u_x \\ u_{xy} & u_{yy} & u_y \\ u_x & u_y & nu \end{vmatrix} = 0 \qquad (n \neq 1),$$
show that u is of the form
$$u = \{x\,\phi(t)+y\,\psi(t)+\chi(t)\}^{n/(n-1)},$$
where
$$0 = x\,\phi'(t)+y\,\psi'(t)+\chi'(t).$$

33. If $u(x,y)$ and $v(x,y)$ satisfy the set of differential equations
$$\begin{Vmatrix} u_{xx} & u_{xy} & u_{yy} \\ v_{xx} & v_{xy} & v_{yy} \end{Vmatrix} = 0,$$
show that they have either the simultaneous forms
$$u = \phi(x)+\psi(y), \qquad v = a\,\phi(x)+a\,\psi(y)+Ax+By+C,$$
where a, A, B, C are constants and ϕ, ψ arbitrary functions; or else the simultaneous forms
$$u = x\,f(t)+y\,g(t)+h(t)$$
$$v = x\int f'(t)k(t)\,dt+y\int g'(t)k(t)\,dt+\int h'(t)k(t)\,dt,$$
where
$$0 = x\,f'(t)+y\,g'(t)+h'(t),$$
and f, g, h, k are arbitrary functions.

34. The Hessian of a homogeneous cubic $u(x,y,z)$ which is the product of linear factors is the cubic itself, save for a constant factor; and, conversely, a homogeneous cubic $u(x,y,z)$ is a product of linear factors, if its Hessian differs from the cubic itself only by a constant factor.

35. (i) Show that every solution of
$$au_x^2+2hu_x u_y+bu_y^2 = 0,$$
where a, b, h are constants, is also a solution of
$$au_{xx}+2hu_{xy}+bu_{yy} = 0,$$
but that the converse is not true.

(ii) Determine the solutions (if any) common to
$$au_x^2+bu_y^2+cu_z^2+2fu_y u_z+2gu_z u_x+2hu_x u_y = 0$$
and
$$au_{xx}+bu_{yy}+cu_{zz}+2fu_{yz}+2gu_{zx}+2hu_{xy} = 0.$$

XIII

DIFFERENTIAL OPERATORS

1. The polynomial differential operator

THE subject-matter of mathematical analysis is, in essence, twofold. It is concerned on the one hand with *numbers* (derived ultimately from the natural numbers), and on the other hand with *operations on numbers* (built up from the four fundamental operations of arithmetic). The peculiar elegance and simplicity of analysis depend largely on an apt use of symbols to represent number, and a similar advantage is to be gained from a convenient symbolic representation of operation. It is, moreover, an important mental economy that symbols of operation should, as far as possible, follow the same laws and processes as symbols of number.

The special contribution of the Differential Calculus to the theory of analytical operations is, of course, that of differentiation, and we have already found it convenient to introduce the 'differential operator' D and subsequently to apply the index notation to this operator, writing $D^n y$ for the nth derivative of y. We now define and discuss the general polynomial operator

$$\phi(D) \equiv \alpha_0 D^n + \alpha_1 D^{n-1} + \ldots + \alpha_{n-1} D + \alpha_n, \tag{1}$$

where $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}, \alpha_n$ may be functions of x, the variable of differentiation. If ξ is any function of x, n times differentiable but otherwise arbitrary, we evidently define

$$\phi(D)\,\xi \equiv \alpha_0 \frac{d^n\xi}{dx^n} + \alpha_1 \frac{d^{n-1}\xi}{dx^{n-1}} + \ldots + \alpha_{n-1}\frac{d\xi}{dx} + \alpha_n\,\xi. \tag{2}$$

It is convenient to include the case in which the coefficients $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ are zero and the operator reduces to α_n. It is clear from (2) that in this degenerate case the effect of the operation is merely to multiply the operand ξ by α_n.

We may combine these polynomial operators by the algebraic operations of *addition*, *subtraction*, and *multiplication*. For addition and subtraction it is sufficient to define

$$\{\phi(D) \pm \psi(D)\}\xi \equiv \phi(D)\xi \pm \psi(D)\xi, \tag{3}$$

or alternatively

$$\{\textstyle\sum \alpha_r D^{n-r}\} \pm \{\sum \beta_r D^{n-r}\} \equiv \sum (\alpha_r \pm \beta_r) D^{n-r}. \tag{4}$$

In virtue of (2) the two definitions are equivalent. For multiplication we define

$$\{\phi(D)\psi(D)\}\xi \equiv \phi(D)\{\psi(D)\xi\}. \tag{5}$$

That is to say, we first perform the operation $\psi(D)$ on ξ, and then, on $\psi(D)\xi$ as new operand, perform the operation $\phi(D)$. Multiplication thus denotes *succession* of operations, the operator nearest to ξ coming first into action, and the order of the succession being therefore from right to left. The definition presupposes that the coefficients in $\psi(D)$ are differentiable at least to the order of $\phi(D)$. We shall make the general assumption that the coefficients in the operators are differentiable as often as we may require.

From (5) and Leibniz's theorem it is clear that the product of any two polynomial operators is also a polynomial operator. From (4) it is clear that the same is true for the sum or difference of two such operators. Hence for the three algebraic operations of addition, subtraction, multiplication the polynomial operators form a closed group. Division, in general, would take us beyond this group, and we exclude it from our present consideration

Now the operations of addition and multiplication defined for arithmetical symbols P, Q, R, .. lead, as we know, to the fundamental laws of arithmetical combination:

$$\begin{aligned} P+Q &= Q+P, \\ P+(Q+R) &= (P+Q)+R, \\ \left.\begin{aligned} P(Q+R) &= PQ+PR \\ (Q+R)P &= QP+RP \end{aligned}\right\}&, \\ P(QR) &= (PQ)R, \\ PQ &= QP. \end{aligned}$$

It is not difficult to verify that the operations of addition and multiplication, as we have defined them for polynomial operators P, Q, R,..., also obey the same laws but with the striking exception of the last of them, $PQ = QP$, the *commutative law of multiplication*. That this is not satisfied in general may be seen from a simple example. Write

$$P = D, \qquad Q = \alpha D. \tag{6}$$

Then $$PQ = \alpha D^2+\alpha' D, \qquad QP = \alpha D^2,$$

and $PQ \neq QP$, unless $\alpha' = 0$, i.e. unless α is a constant.

Two operations are said to be *commutative* or *non-commutative* according as they do or do not satisfy the commutative law of multiplication. Thus differential operators are in general non-commutative. So too we know that in a repeated limit the two operations of passing to the limit are not necessarily commutative, and it is, in fact, because differentiation ultimately involves passage to a limit, that differential operations are

not necessarily commutative. In geometry non-commutative operations are frequent. Thus, in the plane, a pair of inversions is commutative, only if the inverting circles cut orthogonally. In space, rotations are not generally commutative, and a special non-commutative algebra, the theory of quaternions, exists for their analytical representation.

In a non-commutative algebra, then, we may freely use all the machinery of ordinary algebra, except that we must be careful never to disturb the order of factors in a product.

If the operators P, Q are non-commutative, the operator $PQ-QP$ is an important associated operator and is known as their *alternant*.

2. Commutative operators

We now inquire what sort of differential operators are commutative. It is convenient in this connexion to make a notational distinction between constants and polynomial forms with constant coefficients on the one hand and variables and polynomial forms with variable coefficients on the other. For the former we shall use the English 'lower-case' letters (x, y naturally excepted); for the latter the Greek 'lower-case' letters. Capital letters in either alphabet we shall generally reserve for operators (as already D itself). Thus we write

$$P \equiv f(D) \equiv a_0 D^n + a_1 D^{n-1} + \ldots + a_n,$$
$$Q \equiv \phi(D) \equiv \alpha_0 D^n + \alpha_1 D^{n-1} + \ldots + \alpha_n,$$

where the a's are constants and the α's are variables. We can now prove that

(7) *Operators $f(D)$, $g(D)$ with constant coefficients are commutative.*

For suppose

$$f(D) \equiv \sum_{r=0}^{n} a_r D^{n-r}, \qquad g(D) \equiv \sum_{s=0}^{m} b_s D^{m-s}.$$

Then

$$\begin{aligned} f(D)g(D) &= \left\{\sum_{r=0}^{n} a_r D^{n-r}\right\}\left\{\sum_{s=0}^{m} b_s D^{m-s}\right\} \\ &= \sum_{r=0}^{n} \sum_{s=0}^{m} a_r D^{n-r} b_s D^{m-s} \\ &= \sum_{r=0}^{n} \sum_{s=0}^{m} a_r b_s D^{m+n-r-s}, \quad \text{since the } b\text{'s are constants,} \\ &= g(D)f(D), \quad \text{by symmetry.} \end{aligned}$$

More generally, if P is any operator $\phi(D)$, a similar argument shows that $f(P)$, $g(P)$ are commutative. Is this the most general form for a pair of commutative operators?

The answer is in the negative, but it is not easy at this stage to give examples to the contrary. The general theory of such operators leads us to the consideration of Abelian functions, but it is possible to establish some of the important theorems by elementary methods.

Suppose, then, that we have two commutative operators P, Q of orders m, n respectively, and consider the differential equation

$$(P-p)y = 0, \tag{8}$$

where p is an arbitrary constant.

Let $\eta_1,\ldots,\eta_m$ be a linearly distinct set of solutions. Then

$$(P-p)Q\eta_r = Q(P-p)\eta_r = 0.$$

Hence $Q\eta_r$ is also a solution and, by chapter V § 3 (20),

$$Q\eta_r = a_{r1}\eta_1+\ldots+a_{rm}\eta_m \qquad (r = 1,\ldots,m),$$

where the a's are constants.

Thus, if q be a root of the m-ic

$$\begin{vmatrix} a_{11}-q & a_{12} & . & a_{1m} \\ a_{21} & a_{22}-q & . & a_{2m} \\ . & . & . & . \\ a_{m1} & a_{m2} & . & a_{mm}-q \end{vmatrix} = 0, \tag{9}$$

then also

$$\begin{vmatrix} (Q-q)\eta_1 & a_{12} & . & a_{1m} \\ (Q-q)\eta_2 & a_{22}-q & . & a_{2m} \\ . & . & . & . \\ (Q-q)\eta_m & a_{m2} & . & a_{mm}-q \end{vmatrix} = 0,$$

and consequently

$$\begin{vmatrix} \eta_1 & a_{12} & . & a_{1m} \\ \eta_2 & a_{22}-q & . & a_{2m} \\ . & . & . & . \\ \eta_m & a_{m2} & . & a_{mm}-q \end{vmatrix}$$

is a solution of the differential equation

$$(Q-q)y = 0. \tag{10}$$

Since it is linear in the η's, it is also a solution of (8). Hence,

(11) *If P, Q are commutative operators, then, given p, we can find q such that the equations $(P-p)y = 0$, $(Q-q)y = 0$ have a common solution $\eta(p,q)$.*

The constants p, q are therefore connected by some functional relation. A form of this relation is given by (9). It is evidently of degree m in q, but its form in p, which enters through the a_{rs}, is not clear. More directly let us form the eliminant of the differential equations

$$(P-p)y = 0, \qquad (Q-q)y = 0.$$

Their common solution y will be, more generally, a common solution of the $m+n$ differential equations

$$D^r(P-p)y = 0 \quad (r = 0,\dots,n-1), \qquad D^s(Q-q)y = 0 \quad (s = 0,\dots,m-1).$$

These $m+n$ equations are all of order not exceeding $m+n-1$, since the operators P, Q are of respective orders m, n. Accordingly, performing the differentiations D^r, D^s, we may write them as $m+n$ equations linear and homogeneous in the $m+n$ 'unknowns'

$$D^t y \quad (t = 0, 1,\dots, m+n-1).$$

Elimination of these unknowns gives a determinant of order $m+n$ in which p enters to degree n, and q to degree m. Since p, q are constants with one degree of freedom, x must disappear from this eliminant and we are left with an algebraic relation

$$F(p,q) = 0 \tag{12}$$

of degrees n, m in p, q respectively.

Form the differential operator†

$$F(P,Q) \tag{13}$$

and operate on the $\eta(p,q)$ of (11). Then

$$F(P,Q)\eta = F(p,q)\eta = 0.$$

Hence the differential equation

$$F(P,Q)y = 0, \tag{14}$$

which is of order not exceeding $m+n$, is satisfied by every such η. Since p is arbitrary, these are infinitely many, and so (14) is an identity, unless the η's are linearly connected. Suppose a linear relation

$$a_1\eta_1+a_2\eta_2+\dots+a_s\eta_s = 0,$$

where η_r is a solution of $(P-p_r)y = 0$ and the p's are all distinct.‡ Operation with $P, P^2,\dots, P^{s-1}$ on this relation adds the $s-1$ relations

$$a_1p_1\eta_1+a_2p_2\eta_2+\dots+a_sp_s\eta_s = 0,$$
$$\cdots\cdots\cdots\cdots$$
$$a_1p_1^{s-1}\eta_1+a_2p_2^{s-1}\eta_2+\dots+a_sp_s^{s-1}\eta_s = 0.$$

† Since P, Q are commutative, no difficulty arises about the actual form in which $F(P,Q)$ should be written.

‡ We clearly need not allow for more than *one* solution of each distinct equation $(P-p_r)y = 0$, since a sum $a_1\eta_1+a_2\eta_2+\dots$ of solutions of the same equation is itself a solution of the equation.

Since no $a_r \eta_r$ is zero, the determinant of these s equations must vanish. But it can be written as the product of the differences of the p's, which are by hypothesis all distinct. Hence $F(P, Q)$ vanishes identically. Thus

(15) *Any two commutative operators P, Q are connected by an algebraic identity $F(P, Q) = 0$, with constant coefficients.*

We can prove the converse of (11), namely that

(16) *If for every p we can find q such that $(P-p)y = 0$, $(Q-q)y = 0$ have a common solution $\eta(p, q)$, then P, Q are commutative operators.*

For operating with $PQ-QP$ on this η we have

$$(PQ-QP)\eta = Pq\eta - Qp\eta = pq\eta - qp\eta = 0.$$

Hence the equation $$(PQ-QP)y = 0$$

has the infinitely many solutions $\eta(p, q)$, which, we have seen, are linearly distinct. It is thus an identity, and P, Q are commutative.

The converse of (15) is also true, but the proof is rather tedious, since, until P, Q are proved commutative, it is necessary to be careful of the arrangement of P, Q in the polynomial form $F(P, Q)$.

If we interpret $F(P, Q) = 0$ as the equation to a plane curve in coordinates (P, Q), the problem of finding expressions for the operators P, Q is seen to bear analogy with the problem of expressing the coordinates of an algebraic plane curve parametrically. Since this parametric expression is obtained most naturally in terms of Abelian functions, their entry into the theory of commutative operators is accounted for. The special case of the unicursal curve evidently corresponds to the case of a pair of operators expressible as $f(R)$, $g(R)$, i.e. as functions with constant coefficients in some common operator R of lower order.

As an example of a pair of commutative operators not reducible to this form write

$$P \equiv D^2 - 2\theta, \qquad Q \equiv D^3 - 3\theta D - \tfrac{3}{2}\theta', \tag{17}$$

where $$\theta'^2 = 4\theta^3 - g_2\theta - g_3 \tag{18}$$

and the accents denote differentiation. Then

$$D(P-p)-(Q-q) = (\theta-p)D - (\tfrac{1}{2}\theta' - q).$$

But

$$\begin{aligned}[(\theta-p)D-(\tfrac{1}{2}\theta'+q)][(\theta-p)D-(\tfrac{1}{2}\theta'-q)] &= (\theta-p)^2D^2 + \tfrac{1}{4}\theta'^2 - q^2 - \tfrac{1}{2}\theta''(\theta-p) \\ &= (\theta-p)^2(P-p), \quad \text{on reduction from (18),}\end{aligned}$$

if $$4q^2 = 4p^3 - g_2 p - g_3. \tag{19}$$

Thus $\eta(p,q)$ given by the equation

$$\frac{D\eta}{\eta} = \frac{\frac{1}{2}\theta' - q}{\theta - p}$$

is a solution of both $(P-p)y = 0$, $(Q-q)y = 0$. Hence, by (16), P, Q are commutative and by (15), (19) they satisfy the identity

$$4Q^3 = 4P^3 - g_2 P - g_3.$$

Other examples of commutative operators will be given in § 4.

3. Adjoint operators

Associated with the theory of commutative operators is that of *adjoint* operators, which are also of importance in the theory of Differential Equations and elsewhere.

Two polynomial operators P, P' are said to be 'adjoint', if, for every ξ, η sufficiently differentiable,

$$\xi P\eta - \eta P'\xi$$

is an exact derivative, i.e. if

$$\xi P\eta - \eta P'\xi = \text{some } D\{\textstyle\sum \alpha_{rs}\xi^{(r)}\eta^{(s)}\},$$

where $\xi^{(r)}$ denotes $D^r\xi$, etc.

Two different operators P', P'' cannot both be adjoint to P. For, if they were, then

$$\xi P\eta - \eta P'\xi = D\{\textstyle\sum \alpha_{rs}\xi^{(r)}\eta^{(s)}\},$$

$$\xi P\eta - \eta P''\xi = D\{\textstyle\sum \beta_{rs}\xi^{(r)}\eta^{(s)}\},$$

and therefore

$$\eta(P''-P')\xi = D\{\textstyle\sum (\alpha_{rs}-\beta_{rs})\xi^{(r)}\eta^{(s)}\}.$$

But clearly the right-hand side cannot have η as a factor. Hence

(20) *The adjoint, if it exists, is unique.*

Now

$$\xi D^n\eta - \eta(-D)^n\xi = D\{\xi\eta^{(n-1)} - \xi'\eta^{(n-2)} + \ldots (-)^{n-1}\xi^{(n-1)}\eta\}.$$

Writing $\alpha\xi$ for ξ we have also

$$\alpha\xi D^n\eta - \eta(-D)^n\alpha\xi = D\{\ldots\}.$$

Hence

(21) *αD^n and $(-)^n D^n\alpha$ are adjoint.*

Again, more generally,

(22) *If P and P' are adjoint and also Q and Q', then $P \pm Q$ and $P' \pm Q'$ are adjoint, and also PQ and $Q'P'$.*

The first result is obvious. To prove the second we have

$$\xi P\eta - \eta P'\xi = D\{...\}.$$

Replace η by $Q\eta$. Then

$$\xi PQ\eta - (Q\eta)(P'\xi) = D\{...\}.$$

Likewise, since Q and Q' are adjoint,

$$(P'\xi)(Q\eta) - \eta Q'P'\xi = D\{...\}.$$

Hence, by addition,

$$\xi PQ\eta - \eta Q'P'\xi = D\{...\},$$

which proves PQ and $Q'P'$ adjoint.

It follows from (21), (22) that

$$(23)\quad \begin{cases} \alpha_n + \alpha_{n-1}D + \ldots + \alpha_0 D^n \ \textit{and}\ \alpha_n - D\alpha_{n-1} + \ldots (-)^n D^n \alpha_0, \\ \alpha_0 D\alpha_1 D\alpha_2 \ldots D\alpha_n \ \textit{and}\ (-)^n \alpha_n D \ldots \alpha_2 D\alpha_1 D\alpha_0, \\ (D-\alpha_1)(D-\alpha_2)\ldots(D-\alpha_n) \ \textit{and}\ (-)^n (D+\alpha_n)\ldots(D+\alpha_2)(D+\alpha_1) \end{cases}$$

are all pairs of adjoint operators.

By (23), an operator such as

$$\alpha_0 D\alpha_1 D\alpha_2 D\alpha_1 D\alpha_0$$

is identical with its adjoint, and is therefore said to be *self-adjoint*. Such an operator must evidently be of even order. But we may extend the term 'self-adjoint' to such an operator as

$$\alpha_0 D\alpha_1 D\alpha_2 D\alpha_2 D\alpha_1 D\alpha_0$$

that differs from its adjoint only in sign.

By (22), the adjoint of the alternant $PQ-QP$ is the alternant $Q'P'-P'Q'$. Hence, if one alternant vanishes identically, so must the other. Also the adjoint of $F(P.Q) = \sum a_{rs} P^r Q^s$ is $\sum a_{rs} Q'^s P'^r$, i.e. $F(P', Q')$, if P', Q' are commutative. Hence

(24) *If two operators are commutative, their adjoints are also commutative and satisfy the same identity.*

It must be observed that two adjoint operators $\phi(D)$, $\psi(D)$ do not remain adjoint, if the independent variable be changed. For such a change replaces D by some αD, and hence the identity that defines the adjoint becomes

$$\xi\phi(\alpha D)\eta - \eta\psi(\alpha D)\xi = \alpha D\{...\}.$$

Accordingly the adjoint pair is now

$$\alpha^{-1}\phi(\alpha D), \qquad \alpha^{-1}\psi(\alpha D).$$

It can be seen from the second and third forms of (23) that an equation $Py = 0$ is soluble when we know the complete solution of the adjoint

equation $P'y = 0$. We can go farther and state the solutions of one equation explicitly in terms of those of the other. Suppose P is of order m, is written with the coefficient of D^m unity, and has $\eta_1,\ldots,\eta_m$ as a linearly distinct set of solutions. Write

$$\Delta \equiv (D, 1)^{m-1}|\eta_1, \eta_2,\ldots, \eta_m|$$

for the Wronskian of all the solutions, and

$$\Delta_r \equiv (D, 1)^{m-2}|\eta_1,\ldots, \eta_{r-1}, \eta_{r+1},\ldots, \eta_m|$$

for the Wronskian of all the solutions save η_r.

Then we can prove that

(25) Δ_1/Δ, $\Delta_2/\Delta,\ldots$, Δ_m/Δ *form a set of linearly distinct solutions of the adjoint equation* $P'y = 0$.

Now we can show† that the Wronskian of these m functions is just Δ^{-1}, which does not vanish, and so the m functions are linearly distinct. It is sufficient therefore to prove them solutions of $P'y = 0$.

The equation whose solutions are $\eta_2,\ldots, \eta_m$ is

$$R_1 y \equiv \Delta_1^{-1}(D, 1)^{m-1}|y, \eta_2,\ldots, \eta_m| = 0,$$

if we write it so that the coefficient of $D^{m-1}y$ is unity. Then

$$(P-DR_1)y = 0$$

is also an equation of order $m-1$ whose roots are $\eta_2,\ldots, \eta_m$. We must therefore have

$$P-DR_1 = \alpha R_1, \quad \text{for some } \alpha,$$

i.e.
$$P = (D+\alpha)R_1.$$

Hence
$$(D+\alpha)R_1 \eta_1 = P\eta_1 = 0.$$

Now
$$R_1 \eta_1 = \Delta^{-1}(D, 1)^{m-1}|\eta_1, \eta_2,\ldots, \eta_m| = \Delta/\Delta_1.$$

Thus
$$(D+\alpha)(\Delta/\Delta_1) = 0,$$

and therefore
$$(D-\alpha)(\Delta_1/\Delta) = 0.$$

But $P' = -R_1'(D-\alpha)$, since $P = (D+\alpha)R_1$. Thus

$$P'(\Delta_1/\Delta) = 0,$$

and similarly every $P'(\Delta_r/\Delta) = 0$. This proves (25).

If P, Q are commutative, then we have seen that $(P-p)y = 0$, $(Q-q)y = 0$ have a common solution, if $F(p, q) = 0$. This is an equation of degree m in q. Hence, fixing p, we have m corresponding values of q, say $q_1,\ldots, q_m$, in general distinct. Thus, as a set of solutions of $(P-p)y = 0$, we may take $\eta_1,\ldots, \eta_m$, where η_r is the common solution of $(P-p)y = 0$, $(Q-q_r)y = 0$. As we have seen, these will be linearly distinct.

† As in chapter V Exs. 16 (i), 18.

Suppose that the leading terms in P, Q are respectively αD^m, βD^n. Then the coefficient of D^{m+n-1} in the alternant $PQ-QP$ on expansion is

$$m\alpha\beta'-n\alpha'\beta.$$

Since it must vanish identically, we have

$$n\alpha'/\alpha = m\beta'/\beta,$$

and we can therefore find some ξ such that

$$\alpha = \xi^m, \qquad \beta = \xi^n.$$

With ξ as new variable of differentiation, P, Q have leading terms D^m, D^n. We can now prove, in more precise extension of (25), that

(26) *If P, Q are commutative and have leading terms D^m, D^n, and if η_r is now the solution common to*

$$(P-p)y = 0, \qquad (Q-q_r)y = 0 \quad (r = 1, .., m),$$

then Δ_r/Δ is the solution common to the adjoint equations

$$(P'-p)y = 0, \qquad (Q'-q_r)y = 0.$$

As in the proof of (25), define R_1 as the operator, with leading term D^{m-1}, that annihilates $\eta_2,\dots,\eta_m$, so that again

$$P-p = (D+\alpha)R_1, \qquad R_1\eta_1 = \Delta/\Delta_1$$

and therefore

$$(D+\alpha)(\Delta/\Delta_1) = 0, \qquad (D-\alpha)(\Delta_1/\Delta) = 0.$$

Now $$R_1 Q\eta_r = R_1 q_r \eta_r = 0 \qquad (r = 2,\dots, m).$$

Thus $R_1 Q$ annihilates $\eta_2,\dots,\eta_m$, and therefore

$$R_1 Q = Q_1 R_1,$$

where Q_1 is some operator.† Hence

$$(D+\alpha)Q_1 R_1 = (D+\alpha)R_1 Q = (P-p)Q = Q(P-p) = Q(D+\alpha)R_1.$$

Thus $$(D+\alpha)Q_1 = Q(D+\alpha). \tag{27}$$

But, again, $$(Q_1-q_1)R_1\eta_1 = R_1(Q-q_1)\eta_1 = 0,$$

i.e. $$(Q_1-q_1)(\Delta/\Delta_1) = 0,$$

and so, for some operator S_1,

$$(Q_1-q_1) = S_1(D+\alpha).$$

Hence, from (27),

$$\begin{aligned}(Q-q_1)(D+\alpha) &= (D+\alpha)(Q_1-q_1)\\ &= (D+\alpha)S_1(D+\alpha),\end{aligned}$$

and so $$Q-q_1 = (D+\alpha)S_1.$$

† For, otherwise, long division of $R_1 Q$ by R_1 would give as remainder an operator of order $m-2$ annihilating the $m-1$ linearly distinct functions $\eta_2,\dots,\eta_m$.

By taking adjoints we have

$$Q'-q_1 = -S_1'(D-\alpha),$$

and therefore $\qquad (Q'-q_1)(\Delta_1/\Delta) = 0.$

This completes the proof, since we have already shown in (25) that

$$(P'-p)(\Delta_1/\Delta) = 0.$$

4. The operator xD

The operator xD has many important uses and we now establish its elementary properties. They are readily deduced from the corresponding properties of the fundamental operator D.

Since $D^r e^{ax} = a^r e^{ax}$, we have by Leibniz's formula, writing $(n\,!\,r)$ for the binomial coefficients,

$$\begin{aligned} D^r(e^{ax}\xi) &= e^{ax}\{D^r\xi+(n\,!\,1)aD^{r-1}\xi+\ldots+a^r\xi\} \\ &= e^{ax}(D+a)^r\xi. \end{aligned}$$

Multiply by appropriate coefficients α_r and sum for r. We have, then, for any polynomial operator $\phi(D)$, the theorem

(28) $$\phi(D)\,e^{ax}\xi = e^{ax}\,\phi(D+a)\xi,$$

and in particular $\qquad \phi(D)\,e^{ax} = \phi(a)\,e^{ax}.$

Again, since $\qquad D^2\cos ax = -a^2\cos ax,$

we have $\qquad D^{2r}\cos ax = (-a^2)^r\cos ax,$

and similarly for $\sin ax$. Hence by the preceding argument we may write

(29) $$\phi(D^2)(\cos ax,\ \sin ax) = \phi(-a^2)(\cos ax,\ \sin ax).$$

We come now to the operator xD which we shall always denote by δ, the conventions of § 2 notwithstanding. Write

$$x = e^t,$$

and hence $\qquad D_t = xD_x = \delta.$

Thus, from (28), we have

$$\phi(D_t)\,e^{at}\xi = e^{at}\,\phi(D_t+a)\xi,$$

which gives the formula

(30) $$\phi(\delta)\,x^a\xi = x^a\,\phi(\delta+a)\xi,$$

and $\qquad \phi(\delta)\,x^a = \phi(a)\,x^a.$

In particular, if $f_1, f_2, f_3,\ldots$ are polynomials with constant coefficients, then

$$f_1(\delta)\,x^a\xi = x^a f_1(\delta+a)\xi,$$

$$x^b f_1(\delta)\,x^a f_2(\delta)\xi = x^{a+b} f_1(\delta+a)\,f_2(\delta)\xi = x^{a+b} f_2(\delta)\,f_1(\delta+a)\xi,$$

and so on. Hence, generally, we have the theorem

(31) $$x^{a_n} f_n(\delta)\, x^{a_{n-1}} f_{n-1}(\delta) \dots x^{a_1} f_1(\delta)$$
$$= x^{s_n} f_1(\delta)\, f_2(\delta+s_1) \dots f_{n-1}(\delta+s_{n-2}) f_n(\delta+s_{n-1}),$$

where $$s_r = a_1+a_2+\dots+a_r \qquad (r = 1,\dots,n).$$

If $a_r = a$ and $f_r(\delta) \equiv f(\delta)$ $(r = 1,\dots,n)$, we have in particular

(32) $$\{x^a f(\delta)\}^n = x^{na} f(\delta)\, f(\delta+a) \dots f\{\delta+(n-1)a\}.$$

If $a = -1$ and $f(\delta) \equiv \delta$, we have, since $x^{-1}\delta = D$,

(33) $$x^n D^n = \delta(\delta-1)\dots(\delta-n+1).$$

Commutative operators can be freely constructed in terms of the operator δ. Thus write

$$P \equiv x^{-2}\,\delta(\delta-3), \qquad Q \equiv x^{-3}\,\delta(\delta-2)(\delta-4). \tag{34}$$

Then, by (30),

$$PQ = x^{-5}\,\delta(\delta-2)(\delta-3)(\delta-4)(\delta-6) = QP,$$
$$P^3 = x^{-6}\,\delta(\delta-2)(\delta-3)(\delta-4)(\delta-5)(\delta-7) = Q^2.$$

The characteristic (p,q)-relation is thus $p^3 = q^2$, and therefore the equations
$$(P-c^2)y = 0, \qquad (Q-c^3)y = 0$$
have a common solution, which can be shown to be
$$\eta = (\delta-1)e^{cx}.$$

More generally, write

$$\left.\begin{aligned} P &= x^{-m}\,\delta(\delta-n)(\delta-2n)\dots(\delta-mn+n) \\ Q &= x^{-n}\,\delta(\delta-m)(\delta-2m)\dots(\delta-mn+m) \end{aligned}\right\}, \tag{35}$$

where m, n are interprime. We find in similar fashion that
$$PQ = QP, \qquad P^n = Q^m.$$

Thus $P-c^m$, $Q-c^n$ annihilate a common η. It can be shown to be
$$\eta = (\delta-b_1)(\delta-b_2)\dots(\delta-b_g)e^{cx},$$
where $b_1, b_2,\dots, b_g$ are the $\frac{1}{2}(m-1)(n-1)$ numbers less than $(m-1)(n-1)$ that are of the form $rm+sn$, where r, s are positive integers.

Again, if we write

$$\left.\begin{aligned} P = x^{-3}\,\delta(\delta-1)(\delta-5) \qquad Q = x^{-4}\,\delta(\delta-1)(\delta-3)(\delta-6) \\ R = x^{-5}\,\delta(\delta-1)(\delta-3)(\delta-4)(\delta-7) \end{aligned}\right\}, \tag{36}$$

we find that P, Q, R are mutually commutative, and that
$$P^4 = Q^3, \qquad Q^5 = R^4, \qquad R^3 = P^5,$$
$$PR = Q^2, \qquad QR = P^3.$$

We have thus a commutative triad and can contemplate similarly commutative sets of any extent.

5. Generalizations of Leibniz's formula and others

We collect here certain generalizations of Leibniz's formula and of other formulae of the foregoing section.

If
$$\phi(D) \equiv \sum_{r=0}^{n} \alpha_{n-r} D^r,$$
we may write for the 'derived' operator
$$\phi'(D) = \sum_{r=1}^{n} r\alpha_{n-r} D^{r-1},$$
in which the derivation is with regard to D and the coefficients α are accordingly unaffected. We shall interpret $\phi''(D),..., \phi^{(m)}(D)$ in a similar fashion. Now write Leibniz's formula as
$$D^r\xi\eta = \xi D^r\eta + (r\,!\,1)\xi' D^{r-1}\eta + (r\,!\,2)\xi'' D^{r-2}\eta + ... + \xi^{(r)}\eta.$$
Multiply by α_{n-r} and sum for r. We have the formula
$$\phi(D)\,\xi\eta = \xi\,\phi(D)\eta + \xi'\,\phi'(D)\eta + \frac{\xi''}{2!}\phi''(D)\eta + ..., \tag{37}$$
the series on the right terminating, since ϕ is a polynomial.

We may generalize this formula in terms of any two polynomial operators P, Q and their successive alternants. Write
$$Q_1 = PQ - QP, \qquad Q_2 = PQ_1 - Q_1 P, \quad ... .$$
Consider the succession of $n+1$ operators
$$QP^n, \quad PQP^{n-1}, \quad P^2QP^{n-2}, \quad ..., \quad P^nQ. \tag{38}$$
The first order of differences gives
$$Q_1 P^{n-1}, \quad PQ_1 P^{n-2}, \quad ..., \quad P^{n-1}Q_1.$$
Similarly, the second-order differences are
$$Q_2 P^{n-2}, \quad PQ_2 P^{n-3}, \quad ..., \quad P^{n-2}Q_2,$$
and generally the leading difference of order r is $Q_r P^{n-r}$.

Hence, if Δ, E denote, as usual, the difference-operator and the succession-operator of Algebra, where
$$E = 1+\Delta,$$
we may write
$$\Delta^r(QP^n) = Q_r P^{n-r}, \qquad E^n(QP^n) = P^nQ.$$
But
$$E^n = (1+\Delta)^n = 1+(n\,!\,1)\Delta+(n\,!\,2)\Delta^2+...+\Delta^n,$$
where $(n\,!\,r)$ denotes the typical binomial coefficient. Thus we have
$$P^nQ = QP^n + (n\,!\,1)Q_1 P^{n-1} + (n\,!\,2)Q_2 P^{n-2} + ... + Q_n. \tag{39}$$

If we multiply by any *constant* coefficients and sum for n, we have, since the constant coefficients are commutative with Q_r on the right,

$$(40) \qquad f(P)Q = Qf(P)+Q_1 f'(P)+\frac{Q_2 f''(P)}{2!}+\ldots+\frac{Q_m f^{(m)}P}{m!},$$

where m is the order of f.

If we write $P = D$, $Q = \xi$, so that $Q_1 = \xi'$, $Q_2 = \xi''$,..., then we may multiply up in (39) by *variable* coefficients, since they are commutative with the debased operators Q_r. Summing for n, we evidently reproduce the previous formula (37).

If we write the set of operators (38) in the reverse order

$$P^n Q, \quad P^{n-1}QP, \quad P^{n-2}QP^2, \quad \ldots, \quad QP^n,$$

and form the corresponding sets of differences, we get similarly

$$\Delta^r P^n Q = (-)^r P^{n-r} Q_r.$$

The formula $\qquad E^n = 1+(n\,!\,1)\Delta+(n\,!\,2)\Delta^2+\ldots+\Delta^n$

now gives

$$QP^n = P^n Q-(n\,!\,1)P^{n-1}Q_1+(n\,!\,2)P^{n-2}Q_2-\ldots(-)^n Q_n. \qquad (41)$$

If we multiply by constant coefficients and sum for n, we have

$$(42) \quad Qf(P) = f(P)Q-f'(P)Q_1+\frac{1}{2!}f''(P)Q_2-\ldots\frac{(-)^m}{m!}f^{(m)}(P)Q_m.$$

If Q is a mere variable ξ, we may multiply up by variable coefficients in (41) and so get

$$(43) \quad \xi\phi(P)\eta = \phi(P)\xi\eta-\phi'(P)P_1\eta+\frac{1}{2!}\phi''(P)P_2\eta-\ldots\frac{(-)^m}{m!}\phi^{(m)}(P)P_m\eta,$$

where $P_1 = P\xi-\xi P$, $P_2 = PP_1-P_1P$, *etc.*

Finally, if $P = D$, we get a formula in a way reciprocal to (37):

$$(44) \quad \xi\phi(D)\eta = \phi(D)\xi\eta-\phi'(D)\xi'\eta+\frac{1}{2!}\phi''(D)\xi''\eta-\ldots\frac{(-)^m}{m!}\phi^{(m)}(D)\xi^{(m)}\eta.$$

To generalize (28) write

$$D(\xi\eta) = \xi D\eta+\xi'\eta = \xi(D+\xi'/\xi)\eta.$$

Replace η by $(D+\xi'/\xi)\eta$; then

$$D\xi(D+\xi'/\xi)\eta = \xi(D+\xi'/\xi)(D+\xi'/\xi)\eta,$$

i.e. $\qquad D^2(\xi\eta) = \xi(D+\xi'/\xi)^2\eta,$

and generally $\qquad D^n(\xi\eta) = \xi(D+\xi'/\xi)^n\eta.$

If we multiply by variable coefficients and sum for n, we have

$$(45) \qquad \phi(D)\,\xi\eta = \xi\,\phi(D+\xi'/\xi)\eta,$$

and therefore in combination with (37)

$$(46) \quad \phi(D+\xi'/\xi) = \phi(D)+\frac{\xi'}{\xi}\phi'(D)+\frac{\xi''}{\xi}\frac{\phi''(D)}{2!}+\ldots+\frac{\xi^{(m)}}{\xi}\frac{\phi^{(m)}(D)}{m!},$$

where m is the order of ϕ.

To generalize (33) we have from (45)

$$[\phi(D)\,\xi]^2\eta = \phi(D)\,\xi^2\,\phi(D+\xi'/\xi)\eta$$
$$= \xi^2\phi(D+2\xi'/\xi)\,\phi(D+\xi'/\xi)\eta,$$

since, again by (45),

$$\phi(D)\,\xi^2\eta = \xi^2\phi(D+2\xi'/\xi).$$

Also $$[\xi\,\phi(D)]^2\eta = \xi^2\phi(D+\xi'/\xi)\,\phi(D)\eta,$$

if we write $\phi(D)\eta$ for η in (45). By successive application of (45) we thus have

$$(47) \quad \begin{cases} [\phi(D)\,\xi]^n = \xi^n\phi(D+n\xi'/\xi)\ldots\phi(D+2\xi'/\xi)\,\phi(D+\xi'/\xi), \\ [\xi\,\phi(D)]^n = \xi^n\phi\{D+(n-1)\xi'/\xi\}\ldots\phi(D+\xi'/\xi)\phi(D). \end{cases}$$

In the second of these two identities write ξ^{-1} for ξ and then ξD for $\phi(D)$. On multiplying both sides by ξ^n we get

$$(48) \qquad \xi^n D^n = \{\xi D-(n-1)\xi'\}\ldots(\xi D-\xi')\xi D$$

and $$D^n\xi^n = (\xi D+\xi')\ldots\{\xi D+(n-1)\xi'\}(\xi D+n\xi'),$$

the second identity being derived from the first by use of (45).

6. Symbolic operators

It is natural to inquire how far the foregoing arguments apply to operators that are no longer polynomial. Inverse operators, of which D^{-1} is the simplest, evidently require the Integral Calculus for their proper development and are consequently beyond our present scope.

On the other hand, infinite direct operators $\sum^{\infty} \alpha_r D^r$ give rise to difficulties of another order. The meaning to be attached to $\sum^{\infty} \alpha_r D^r\xi$ depends on its convergence, and this must depend on the nature of ξ. We cannot therefore discuss the convergence of the operator by itself apart from knowledge of the operand. For instance,

$$(1-D)\sum_{r=0}^{n-1} D^r\xi = \xi-D^n\xi,$$

and thus $$\sum_{r=0}^{\sigma} D^r\xi$$

is a solution of the differential equation

$$(1-D)y = \xi,$$

if
$$\lim_{n\to\infty} D^n\xi = 0,$$
i.e. certainly, if $\sum D^r\xi$ converges.

But there are other difficulties. Consider Taylor's formula

$$f(x+h) = f(x)+hf'(x)+\frac{h^2f''(x)}{2!}+\dots \text{ to } \infty$$
$$= \left(1+hD+\frac{h^2D^2}{2!}+\dots \text{ to } \infty\right)f(x)$$
$$= \exp(hD)f(x), \quad \text{formally.}$$

Thus, *on a function satisfying Taylor's formula*, the operator $\exp(hD)$ changes x into $x+h$. But, as we have already seen in chapter VIII, the conditions under which Taylor's formula is valid are far from simple, mere convergence of the series on the right being insufficient to secure validity.

For reasons of this kind it is desirable to regard all such infinite operators as purely symbolic and to treat the results to which they lead as hypothetical results to be established, if possible, by other arguments. Such operators are none the less valuable and potent instruments of research, and may indicate results otherwise unsuspected. But we must remember that these results form the starting-point and not the culmination of rigorous analysis.

In the restricted algebraic field in which the operand is merely a polynomial in x these infinite operators and their results are above suspicion. For then the operators are only nominally infinite; since, if the operand is of the nth degree in x, the operators effectively terminate at D^n. In this case we can say strictly that $\exp(hD)$ is the operator that changes x into $x+h$, and hence we can identify $\exp(D)$ with the algebraic succession-operator E that changes x into $x+1$. If Δ is the corresponding difference-operator, we have in this case rigorously

$$\exp(D) = 1+\Delta. \tag{49}$$

Then, taking logarithms and expanding, we may write, at least symbolically,

$$D = \Delta-\tfrac{1}{2}\Delta^2+\tfrac{1}{3}\Delta^3-\dots. \tag{50}$$

But, if the operand is of degree n in x, its differences after those of the nth order vanish, and the series on the right therefore terminates at Δ^n. In this case we can establish (50) rigorously, for, if $f(x)$ is such an operand and p is any positive integer, then

$$f(x+p) = (1+\Delta)^p f(x)$$

gives

$$\left(1+pD+\frac{p^2D^2}{2!}+\dots+\frac{p^nD^n}{n!}\right)f(x)$$
$$=\{1+p\Delta+\tfrac{1}{2}p(p-1)\Delta^2+\tfrac{1}{3}p(p-1)(\tfrac{1}{2}p-1)\Delta^3+\dots+(p!n)\Delta^n\}f(x)$$

This, regarded as an equation in p, is satisfied by all positive integer values of p and is therefore an identity. We may accordingly equate like powers of p on the two sides. The coefficients of p itself give (50). The coefficients of p^2 give

$$\tfrac{1}{2}D^2=\tfrac{1}{2}\Delta^2-\tfrac{1}{3}(1+\tfrac{1}{2})\Delta^3+\tfrac{1}{4}(1+\tfrac{1}{2}+\tfrac{1}{3})\Delta^4-\dots, \tag{51}$$

and so on.

It is of interest to test the identities of operators (50), (51) on functions no longer polynomial. On $(1+a)^x$ we obtain from (50), (51)

$$D(1+a)^x=(a-\tfrac{1}{2}a^2+\tfrac{1}{3}a^3-\dots\text{ to }\infty)(1+a)^x=(1+a)^x\log(1+a),$$
$$\tfrac{1}{2}D^2(1+a)^x=\{\tfrac{1}{2}a^2-\tfrac{1}{3}(1+\tfrac{1}{2})a^3+\dots\text{ to }\infty\}(1+a)^x=(1+a)^x\{\log(1+a)\}^2$$

if $|a|<1$, which are true.

On x^{-1} we obtain from (50)

$$\frac{1}{x^2}=\frac{1}{x(x+1)}+\frac{1}{x(x+1)(x+2)}+\frac{2!}{x(x+1)(x+2)(x+3)}+\dots\text{ to }\infty \tag{52}$$

The remainder after n terms is found to be

$$\frac{n!}{x^2(x+1)(x+2)\dots(x+n)}=x^{-2}\Big/\prod_{r=1}^{n}\left(1+\frac{x}{r}\right)$$

The denominator diverges to infinity or to zero according as x is positive or negative, and therefore (52) is true only if x is positive.

Finally, since $\Delta\sin 2\pi x=0$, the identities (50), (51) are completely invalid except at isolated values of x.

7. Partial differential operators

We can extend our theory of polynomial operators to partial differential operators. In a field of n independent variables $(x_1, x_2, \dots, x_n)$ write, as usual, ∂_r for $\partial/\partial x_r$. Then the general polynomial operator in this field is

$$\phi(\partial_1,\partial_2,\dots,\partial_n)=\sum\alpha\,\partial_1^a\partial_2^b\dots\partial_n^k$$

Addition, subtraction, and multiplication can be defined for these operators as for the ordinary operator. No new principle is involved, and the partial operators are seen without difficulty to satisfy the fundamental arithmetical laws, the commutative law of multiplication, of course, excepted, since the partial operator includes the ordinary operator as the special case $n=1$. The operand is to be supposed

differentiable (in the extended sense of chapter VI § 4) as often as may be required, so that, in particular, for every r, s,

$$\frac{\partial^2 \xi}{\partial x_r \partial x_s} = \frac{\partial^2 \xi}{\partial x_s \partial x_r},$$

and the fundamental operators $\partial_1, \partial_2, \ldots, \partial_n$ are therefore all mutually commutative. As a consequence

(53) *Two operators $f(\partial_1, \partial_2, \ldots, \partial_n)$ and $g(\partial_1, \partial_2, \ldots, \partial_n)$ with constant coefficients are commutative,*

the proof being similar to that for ordinary operators.

Again, on a sufficiently differentiable operand $f(\xi_1, \xi_2, \ldots, \xi_p)$, we have for every x_s

$$\frac{\partial f}{\partial x_s} = \frac{\partial \xi_1}{\partial x_s}\frac{\partial f}{\partial \xi_1} + \frac{\partial \xi_2}{\partial x_s}\frac{\partial f}{\partial \xi_2} + \ldots + \frac{\partial \xi_p}{\partial x_s}\frac{\partial f}{\partial \xi_p}.$$

Multiply by α_s and sum for s. Then, if P is the linear operator $\sum \alpha_s \partial_s$,

$$(54) \qquad Pf = (P\xi_1)\frac{\partial f}{\partial \xi_1} + (P\xi_2)\frac{\partial f}{\partial \xi_2} + \ldots + (P\xi_p)\frac{\partial f}{\partial \xi_p},$$

i.e. the formula of the total differential extends to any linear operator P.

The general theory of commutative partial differential operators has yet to be worked out. The case of commutative linear operators is covered by the following theorem.

(55) *Given a set of p commutative, linearly independent linear operators in a field of n variables ($p < n$), we can transform to new variables $(\xi_1, \xi_2, \ldots, \xi_n)$ such that the operators become the fundamental operators*

$$\frac{\partial}{\partial \xi_1}, \quad \frac{\partial}{\partial \xi_2}, \quad \ldots \quad \frac{\partial}{\partial \xi_p}.$$

Let P, Q be any two of the commutative operators and η a function annihilated by P. Then

$$PQ\eta = QP\eta = 0.$$

Thus $Q\eta$ is also annihilated by P. But the general solution of the partial differential equation $Py = 0$ is $y = f(\eta_2, \ldots, \eta_n)$, where $\eta_2, \ldots, \eta_n$ are $n-1$ functionally independent solutions. We are thus led to the $n-1$ equations

$$Q\eta_r = f_r(\eta_2, \ldots, \eta_n)\dagger \qquad (r = 2, \ldots, n). \tag{56}$$

By (54), Q, operating on any function of the η's only, can be written

$$(Q\eta_2)\frac{\partial}{\partial \eta_2} + \ldots + (Q\eta_n)\frac{\partial}{\partial \eta_n},$$

† Here f_r denotes any independent function and not, as usually, a derived functional form.

and thus, in virtue of equations (56), is completely expressed in the field of $n-1$ variables $(\eta_2,\ldots,\eta_n)$. In this field the differential equation $Qy = 0$ has at most $n-2$ functionally independent solutions, say

$$\zeta_3(\eta_2,\ldots,\eta_n),\quad \ldots,\quad \zeta_n(\eta_2,\ldots,\eta_n).$$

Since these are functions of the η's, they are also annihilated by P. They constitute, in fact, a complete set of $n-2$ functionally independent common solutions of the equations $Py = 0 = Qy$.

If R be another operator of the commutative set, we prove in similar fashion that $R\zeta$ is also a common solution of the two equations, that R may be completely expressed in the field of $n-2$ variables $(\zeta_3,\ldots,\zeta_n)$, and that we may hence determine a complete set of $n-3$ functionally independent common solutions of $Py = 0$, $Qy = 0$, $Ry = 0$.

Now let the p commutative operators be $P_1, P_2,\ldots, P_p$. Proceeding in the above manner we can find

$$\xi_{p+1}, \xi_{p+2},\ldots, \xi_n$$

a complete set of $n-p$ common solutions of the equations

$$P_1 y = 0,\qquad P_2 y = 0,\quad \ldots,\quad P_p y = 0.$$

Again, by expressing P_1 in the field of $n-p+1$ common solutions of $P_2 y = 0,\ldots, P_p y = 0$, we can similarly obtain a common solution of the equations

$$P_1 y = 1,\qquad P_2 y = 0,\quad \ldots,\quad P_p y = 0.$$

Call this ξ_1 and generally let

$$\xi_r \quad (r = 1,\ldots,p)$$

be a common solution of the equations

$$P_r y = 1,\qquad P_s y = 0 \quad (s = 1,\ldots,r-1, r+1,\ldots,p).$$

Then in the field of variables $(\xi_1, \xi_2,\ldots, \xi_n)$

$$P_r = (P_r\xi_1)\frac{\partial}{\partial\xi_1} + (P_r\xi_2)\frac{\partial}{\partial\xi_2} + \ldots + (P_r\xi_n)\frac{\partial}{\partial\xi_n}.$$

But P_r annihilates every ξ except ξ_r, and $P_r\xi_r = 1$. Hence

$$P_r = \frac{\partial}{\partial\xi_r} \quad (r = 1, 2,\ldots,p),$$

and the theorem is proved.

It follows from (55) that, in a field of n variables, given any set of p $(p \leqslant n)$ commutative linear operators, we can always find $n-p$ other operators sufficient to form a complete set of n commutative linear operators. For these remaining operators are $\partial/\partial\xi_{p+1},\ldots, \partial/\partial\xi_n$ in the transformed system.

The condition that the p operators be linearly independent is necessary to the proof, for, if P_1, say, is expressible linearly in terms of $P_2,\dots,P_p$, it vanishes identically when expressed in a field of common solutions of $P_2 y = 0,\ \dots,\ P_p y = 0$, and the equations $P_1 y = 1$, $P_2 y = 0,\dots, P_p y = 0$ are incompatible and have no common solution ξ_1.

To cover the more general case suppose then, that we have a set of $p+m$ commutative linear operators P_r of which $P_1,\dots,P_p$ are linearly independent and the others $P_{p+1},\dots,P_{p+m}$ expressible linearly in terms of them. By the theorem (55) itself we can transform to a field of variables $(\xi_1,\dots,\xi_n)$ in which the p operators $P_1,\dots,P_p$ become

$$\frac{\partial}{\partial\xi_1},\quad \dots,\quad \frac{\partial}{\partial\xi_p}.$$

By hypothesis, any other operator of the set is then expressible as

$$P_{p+r} = \alpha_{r1}\frac{\partial}{\partial\xi_1}+\dots+\alpha_{rp}\frac{\partial}{\partial\xi_p}\quad (r = 1,\dots,m).$$

These operators are evidently commutative with the operators $\partial/\partial\xi_s$ $(s = 1,\dots,p)$, if and only if every coefficient α_{rs} is annihilated by every $\partial/\partial\xi_s$, i.e. if and only if every α_{rs} is a function of variables $\xi_{p+1},\dots,\xi_n$ alone. In this case, too, the m operators P_{p+r} are also commutative amongst themselves. Thus

(57) *A set of $p+m$ commutative linear operators of which only p are linearly independent can, by transformation to suitable variables $(\xi_1,\dots,\xi_n)$, be expressed in the forms*

$$\frac{\partial}{\partial\xi_r}\quad (r = 1,\dots,p),\qquad \sum_{r=1}^{p}\alpha_{rs}\frac{\partial}{\partial\xi_r}\quad (s = p+1,\dots,p+m)$$

where every α_{rs} is a function of $\xi_{p+1},\dots,\xi_n$ alone.

In a field of n variables there cannot, of course, be more than n linearly distinct linear operators. If $p = n$, it is a corollary of (57) that the $m+n$ commutative operators can be transformed into

$$\frac{\partial}{\partial\xi_r}\quad (r = 1,\dots,n),\qquad \sum_{r=1}^{n}a_{rs}\frac{\partial}{\partial\xi_r}\quad (s = n+1,\dots,n+m),$$

where every a_{rs} is a constant.

8. Alternants of linear operators

For non-commutative linear operators we are led to consider the alternant. We must first observe that

(58) *The alternant of two linear operators is itself a linear operator.*

For, if P, Q are $\sum \alpha_r \partial_r$, $\sum \beta_s \partial_s$, then

$$PQ = \sum_r \sum_s \alpha_r \beta_s \partial_r \partial_s + \sum_s \sum_r \alpha_r (\partial_r \beta_s) \partial_s,$$

$$QP = \sum_r \sum_s \alpha_r \beta_s \partial_r \partial_s + \sum_r \sum_s \beta_s (\partial_s \alpha_r) \partial_r,$$

and the alternant is the linear operator

$$\sum_s \sum_r \alpha_r (\partial_r \beta_s) \partial_s - \sum_r \sum_s \beta_s (\partial_s \alpha_r) \partial_r.$$

It is most easily written down, if we operate with P on the *coefficients* of Q, with Q on the *coefficients* of P, and take the difference.

Let us denote the alternant $PQ-QP$ by the symbol (P, Q), the order of the letters being relevant. Then since (P, Q) is linear, we can form its alternants with P and with Q respectively, namely, $\{P, (P, Q)\}$ and $\{Q, (P, Q)\}$. These are again linear operators and may be used to form other alternants and so on. It may be remarked that given any three linear operators P, Q, R we have identically

$$(59) \qquad \{P, (Q, R)\} + \{Q, (R, P)\} + \{R, (P, Q)\} = 0.$$

For, written at length, the left-hand side of (59) is

$$P(QR-RQ)-(QR-RQ)P+Q(RP-PR)-(RP-PR)Q+$$
$$+R(PQ-QP)-(PQ-QP)R,$$

and the operators cancel in pairs.

In particular, if P, Q are commutative, we deduce that

$$(60) \qquad \{P, (Q, R)\} = \{Q, (P, R)\}.$$

Now the series of repeated alternants that can be constructed from a given pair of linear operators, in general, does not terminate. Consider, for instance, the pair

$$P = \sum_r \partial_r, \qquad Q = \sum_r x_r^k \partial_r.$$

The successive alternants are, except for numerical multipliers,

$$(P, Q) = \sum_r x_r^{k-1} \partial_r,$$

$$\{P, (P, Q)\} = \sum_r x_r^{k-2} \partial_r,$$

$$\{(P, Q), Q\} = \sum_r x_r^{2k-2} \partial_r, \quad \text{etc.}$$

A typical member of the series is

$$\sum_r x_r^{pk-p-q} \partial_r,$$

where p, q are any positive integers. The exponent $pk-p-q$ gives rise to a double sequence in (p, q), all the terms of which are distinct, if

k is irrational. The alternants in such a case are therefore infinitely numerous.

Now, if η is a function annihilated by two linear operators P, Q, it is annihilated also by their alternant (P, Q) and so by every repeated alternant. But, if η is annihilated by n linearly independent linear operators in a field of n variables, we have n independent equations of the form

$$\left.\begin{array}{l} \alpha_{11}\partial_1\eta+\alpha_{12}\partial_2\eta+\ldots+\alpha_{1n}\partial_n\eta=0 \\ \cdots\cdots\cdots \\ \alpha_{n1}\partial_1\eta+\alpha_{n2}\partial_2\eta+\ldots+\alpha_{nn}\partial_n\eta=0 \end{array}\right\},$$

which are satisfied only by $\partial_1\eta=\partial_2\eta=\ldots=\partial_n\eta=0$, giving the trivial solution $\eta=$ constant. Thus if P, Q are linear operators, the pair of differential equations $Py=0$, $Qy=0$ has no common solution other than $y=$ constant unless the successive alternants of P, Q form a closed group of less than n linearly independent operators. Evidently a similar condition applies to the successive alternants of $P_1, \ldots, P_m$ if the m equations $P_r y=0$ $(r=1, \ldots, m)$ are to have a common solution other than $y=$ constant. We are thus led to consider the possibility of a group of operators $P_1, \ldots, P_p$ such that, for every pair of suffixes r, s,

$$P_rP_s-P_sP_r=\alpha_{rs1}P_1+\ldots+\alpha_{rsp}P_p. \tag{61}$$

It can be seen that we obtain equations of the same form, if we replace any P by a linear combination

$$\lambda_1P_1+\ldots+\lambda_pP_p$$

of the P's. Hence too we must suppose the P's linearly distinct, for otherwise some of the $\frac{1}{2}p(p-1)$ equations are redundant.

We shall see that in virtue of the equations (61) we can determine p linear combinations of the P's that are strictly commutative and are therefore expressible as $\partial/\partial\xi_1, \ldots, \partial/\partial\xi_p$. To do this solve the P's for any p of the fundamental operators, say $\partial/\partial x_1, \ldots, \partial/\partial x_p$, and so obtain p relations of the form

$$\partial_r=\sum_{s=1}^{p}\lambda_{rs}P_s-\sum_{s=p+1}^{n}\alpha_{rs}\partial_s \qquad (r=1, \ldots, p).$$

Writing
$$Q_r=\sum_{s=1}^{p}\lambda_{rs}P_s \qquad (r=1, \ldots, p)$$

we have
$$Q_r=\partial_r+\sum_{s=p+1}^{n}\alpha_{rs}\partial_s.$$

The Q's, since they are linear combinations of the P's, obey relations similar to (61), say

$$Q_rQ_s-Q_sQ_r=\sum_{t=1}^{p}\mu_tQ_t. \tag{62}$$

Now the alternants of the Q's are independent of $\partial_1,\ldots,\partial_p$, since these have constant coefficients in the Q's. But on expansion we have

$$\sum \mu_t Q_t = \mu_1 \partial_1 + \ldots + \mu_p \partial_p + \sum_{s=p+1}^{n} \mu_t \alpha_{ts} \partial_s.$$

Hence we must have in every equation (62)

$$\mu_r = 0 \qquad (r = 1,\ldots,p).$$

In other words, the Q's are commutative and they are linear combinations of the P's. Hence, by (55), the operators Q may be transformed by appropriate change of variables into $\partial/\partial\xi_1,\ldots,\partial/\partial\xi_p$ and accordingly the operators P into linear combinations of them. Thus

(63) *If the linear operators $P_1,\ldots,P_p$ are such that every alternant is expressible as $\sum \lambda_r P_r$, then we can transform to variables $(\xi_1,\ldots,\xi_n)$ such that*

$$P_r = \alpha_{r1}\frac{\partial}{\partial \xi_1} + \ldots + \alpha_{rp}\frac{\partial}{\partial \xi_p} \quad (r = 1,\ldots,p),$$

where α_{rs} is any function of the variables and, to avoid triviality, $p < n$.

If the P's are not all linearly independent, the theorem can be applied to a linearly independent set of them of which all the others are linear functions. This set, and hence all the P's, are then expressible in terms of some $\partial/\partial\xi_1,\ldots,\partial/\partial\xi_m$ where $m < p$, and so the theorem still applies but with some of the α's identically zero.

As a corollary to (63) we have

(64) *If in a field of n variables the operators $P_1,\ldots,P_r$ together with their alternants form a set of p linearly independent operators, then the equations $P_1 y = 0,\ldots, P_r y = 0$ have just $n-p$ functionally independent solutions.*

For, in the notation of (63), the variables $\xi_{p+1},\ldots,\xi_n$ that do not enter explicitly into the operators constitute such a set of solutions.

9. Adjoint operators

Returning now to operators of any order, we may generalize for a field of n variables the definition of adjoint operators of § 3. We say that two operators P, P' are adjoint, if, given any ξ, η sufficiently differentiable,

$$\xi P\eta - \eta P'\xi = \sum_{r=1}^{p} \partial_r \phi_r(\xi, \eta, \ldots, \partial_s \xi, \partial_s \eta, \ldots),$$

where ϕ_r is a function of ξ, η and their partial derivatives. It is evident, as for ordinary operators, that adjoint operators do not necessarily remain adjoint for transformation of the independent variable. From this definition we can deduce, as for ordinary operators, that

(65) (i) $\sum \alpha \partial_1^a \partial_2^b \ldots \partial_n^k$ *and* $\sum (-)^{a+b+\ldots+k} \partial_1^a \partial_2^b \ldots \partial_n^k \alpha$ *are adjoint;*

(ii) *given P, there is one and only one adjoint P';*

(iii) *if Q and Q' are also adjoint, then $P \pm Q$ and $P' \pm Q'$ are adjoint and so are PQ and $Q'P'$;*

(iv) *if P and Q are commutative, so are their adjoints.*

In the absence of a theory of the general solution of a partial differential equation of any order it is clearly fruitless to look for an analogue of (25) in the case of adjoint partial operators.

10. Exact derivatives

We conclude with the consideration of a differential operator of interest in the Calculus of Variations. Suppose y an arbitrary function of a single variable x, and, writing y_n for the nth derivative $D^n y$, consider functions of the type $f(x, y, y_1, \ldots, y_r, \ldots)$, where f is a polynomial in every argument, save possibly x. Differentiation of such a function leads to a function of the same type with the highest suffix increased by unity, and we have, in fact, the identity of operators

$$D = \frac{\partial}{\partial x} + y_1 \frac{\partial}{\partial y} + y_2 \frac{\partial}{\partial y_1} + \ldots + y_{r+1} \frac{\partial}{\partial y_r} + \ldots,$$

or more concisely

$$D = \partial_x + y_1 \partial + y_2 \partial_1 + \ldots + y_{r+1} \partial_r + \ldots. \tag{66}$$

If in this way we have

$$Df(x, y, y_1, \ldots, y_{n-1}) = F(x, y, y_1, \ldots, y_{n-1}, y_n), \tag{67}$$

the form F is said to be an *exact derivative*. Clearly, not every function of the type f is an exact derivative, as, for instance,

$$y_2^3, \qquad xy_1 + 2y, \qquad y_n \sin x,$$

and we are thus led to consider criteria of exact derivatives. The following result is fundamental:

(68) *Exact derivatives are annihilated by the operator*

$$\Delta = \partial - D\partial_1 + D^2 \partial_2 - \ldots (-)^r D^r \partial_r \ldots.$$

The operator may be written as formally infinite, since it must cease at ∂_n, if y_n is the highest derivative in the operand. This will be the case, if the operand is the exact derivative (67), where y_{n-1} is the highest derivative in the primitive function f. Consider now the alternants of

D with the operators ∂_r, themselves mutually commutative. We have from (66)

$$\begin{aligned} \partial D &= D\partial, \\ \partial_1 D &= D\partial_1+\partial, \\ \partial_2 D &= D\partial_2+\partial_1, \\ &\cdots \\ \partial_n D &= D\partial_n+\partial_{n-1}. \end{aligned}$$

Multiply these equations in succession by 1, $-D$, $D^2,\ldots,(-D)^n$ and add. We get after cancellation

$$\Delta D = (-)^n D^{n+1}\partial_n,$$

and hence, on (67),
$$\Delta F = (-)^n D^{n+1}\partial_n f = 0,$$
since y_n does not appear in f.

This proves the condition necessary. Conversely, we can prove that

(69) *Functions $F(x,y,y_1,\ldots,y_n)$ annihilated by Δ are exact derivatives.*

First suppose that F is homogeneous of degree i in its arguments $y, y_1,\ldots,y_n$, so that, by Euler's theorem,†

$$(y\partial+y_1\partial_1+\ldots+y_n\partial_n)F = iF. \tag{70}$$

Introduce the set of operators

$$\Delta_r = \partial_r - D\partial_{r+1}+\ldots(-)^{n-r}D^{n-r}\partial_n \qquad (r=1,\ldots,n).$$

We have at once
$$\begin{aligned} \Delta+D\Delta_1 &= \partial, \\ \Delta_1+D\Delta_2 &= \partial_1, \\ &\cdots \\ \Delta_n &= \partial_n. \end{aligned}$$

Hence, in (70),

$$\begin{aligned} iF &= y(\Delta+D\Delta_1)F+y_1(\Delta_1+D\Delta_2)F+\ldots+y_n\Delta_n F \\ &= y\Delta F+D(y\Delta_1 F)+D(y_1\Delta_2 F)+\ldots+D(y_{n-1}\Delta_n F), \end{aligned}$$

and therefore, if F is annihilated by Δ, it is expressible as the exact derivative
$$F = i^{-1}D(y\Delta_1+y_1\Delta_2+\ldots+y_{n-1}\Delta_n)F.$$

If F is not homogeneous, separate it into homogeneous parts $F_1, F_2,\ldots, F_s$, so that
$$\Delta F_1+\Delta F_2+\ldots+\Delta F_s = 0.$$

Now, by (66), the operator D is homogeneous in $y,\ldots,y_n$ of degree zero, and so the operator Δ is homogeneous in them of degree -1. Hence, since $F_1,\ldots,F_s$ are homogeneous of distinct degrees, so also are $\Delta F_1,\ldots,\Delta F_s$, and thus their sum cannot vanish identically, unless each

† Chapter VI (47).

separately vanish. By the foregoing argument each F_r is then an exact derivative, and accordingly every F annihilated by Δ is an exact derivative.

It is to be remarked that, when F is an exact derivative, the determination of the primitive function f has been accomplished by direct differentiation without need to employ the Integral Calculus.†

WORKED EXAMPLE

If X is any solution of the differential equation

$$(\delta-a)f(\delta)X=kx^mX,$$

and n any positive integer, prove that

$$\{x^{-m}(\delta-a)f(\delta)-k\}^n x^{-m}f(\delta)X=(-m)^n n!\{x^{-m}f(\delta)\}^{n+1}X.$$

In particular, if m, n are positive integers, prove that

$$(D^m-a^m)^n\frac{e^{ax}}{x}=(-m)^n n!(x^{-1}D^{m-1})^n\frac{e^{ax}}{x}.$$

Write
$$\xi_n=x^{-mn}\prod_{r=0}^{n-1}f(\delta-rm)X.$$

Then
$$\begin{aligned}(\delta+mn-a)f(\delta)\xi_n&=x^{-mn}(\delta-a)\prod_{r=0}^{n}f(\delta-rm)X\\&=x^{-mn}\prod_{r=1}^{n}f(\delta-rm)\,(\delta-a)f(\delta)X\\&=x^{-mn}\prod_{r=1}^{n}f(\delta-rm)kx^mX,\quad\text{by definition of } X,\\&=kx^m.\,x^{-mn}\prod_{r=0}^{n-1}f(\delta-rm)X\\&=kx^m\xi_n.\end{aligned}$$

Thus
$$\begin{aligned}\{x^{-m}(\delta-a)f(\delta)-k\}\xi_n&=-mn\,x^{-m}f(\delta)\xi_n\\&=-mn\,x^{-m}f(\delta)\,x^{-mn}\prod_{r=0}^{n-1}f(\delta-rm)X\\&=-mn\,x^{-m(n+1)}\prod_{r=0}^{n}f(\delta-rm)X\\&=-mn\,\xi_{n+1}.\end{aligned}$$

Using this as a recurrence formula and remembering that $\xi_1=x^{-m}f(\delta)X$, we obtain
$$\begin{aligned}\{x^{-m}(\delta-a)f(\delta)-k\}^n x^{-m}f(\delta)X&=(-m)^n n!\,\xi_{n+1}\\&=(-m)^n n!\{x^{-m}f(\delta)\}^{n+1}X,\end{aligned}$$

since
$$\{x^{-m}f(\delta)\}^n=x^{-mn}\prod_{r=0}^{n-1}f(\delta-rm),\quad\text{by (32).}$$

For the second part take

$$f(\delta)=\delta(\delta-1)\ldots(\delta-m+2),\qquad a=m-1,\qquad k=a^m.$$

† The subject is further discussed in Elliott, *Messenger of Math.* 43 (1913), 87–92; 45 (1915), 33–9; Chaundy, *ibid.* 45 (1916), 168–76.

Then X satisfies the differential equation

$$\delta(\delta-1)\ldots(\delta-m+1)X = a^m x^m X,$$

i.e.

$$D^m X = a^m X,$$

so that we may take X to be e^{ax}. Thus

$$x^{-m}(\delta-a)f(\delta) = D^m, \qquad x^{-m}f(\delta) = x^{-1}D^{m-1},$$

and the result becomes

$$(D^m-a^m)^n x^{-1}D^{m-1}e^{ax} = (-m)^n n!(x^{-1}D^{m-1})^{n+1}e^{ax}.$$

Since $x^{-1}D^{m-1}e^{ax} = a^{m-1}e^{ax}/x$, we may write this more simply as

$$(D^m-a^m)^n\frac{e^{ax}}{x} = (-m)^n n!(x^{-1}D^{m-1})^n\frac{e^{ax}}{x},$$

which is the second result.

EXAMPLES XIII

1. If $y(x)$ is any differentiable function, prove that

$$y^2D^3y^2 = 2D(y^3D^2y).$$

2. If n is odd, prove that

$$(D^2+3^2)(D^2+5^2)\ldots(D^2+n^2)\sin^n x = n!\sin x,$$
$$(D^2+1)(D^2+3^2)\ldots\{D^2+(n-2)^2\}\sin^n x = (n-1)!\sin nx.$$

3. For any positive integer r, prove that

$$f(\delta)D^r = D^r\left\{f(\delta)-rf'(\delta)+\frac{1}{2!}r^2f''(\delta)-\ldots\right\},$$

$$D^r f(\delta) = \left\{f(\delta)+rf'(\delta)+\frac{1}{2!}r^2f''(\delta)+\ldots\right\}D^r,$$

and more generally, if $g(x)$ is a polynomial and we write $\delta^s g(x) = g_s(x)$, prove that

$$f(\delta)\,g(D) = g(D)\,f(\delta)-g_1(D)\,f'(\delta)+\frac{1}{2!}g_2(D)\,f''(\delta)-\ldots,$$

$$g(D)\,f(\delta) = f(\delta)\,g(D)+f'(\delta)\,g_1(D)+\frac{1}{2!}f''(\delta)\,g_2(D)+\ldots.$$

4. If $D_1 = x^{\frac{1}{2}}D$, $D_2 = x^{-\frac{1}{2}}D$, prove that

$$x(D_2D_1-D_1D_2), \qquad D_2D_1+D_1D_2$$

are commutative, and that

$$x^2(D_2D_1-D_1D_2), \qquad x^2(D_2D_1+D_1D_2), \qquad x^4(D_2D_1^2D_2-D_1D_2^2D_1)$$

are commutative.

5. If m, n, p are positive integers, prove that

(i) $D^n x^{n-1}\log x = (n-1)!\,x^{-1}$,

(ii) $D^n x^n\log x = n!\left(\log x+1+\dfrac{1}{2}+\ldots+\dfrac{1}{n}\right)$,

(iii) $x^{n+1}D^n x^{-1}\log x = (-)^n n!\left(\log x-1-\dfrac{1}{2}-\ldots-\dfrac{1}{n}\right)$,

(iv) $D^m(xD-m)^n x^m(\log x)^n = m!\,n!$,

(v) $x^{m-n}D^m x^n(\log x)^p$ is the coefficient of t^p in the expansion of

$$p!(t+n)(t+n-1)\ldots(t+n-m+1)x^t.$$

6. Show that

(i) $D^n x^{n-1}\exp(x^{-1}) = (-)^n x^{-(n+1)}\exp(x^{-1})$,

(ii) $x^{n+1}D^n x^{n-1}(1+x^{-1})^{-a} = a(a+1)\dots(a+n-1)(1+x^{-1})^{-(a+n)}$.

7. If $xX = 1$ and $y(x)$ is any differentiable function, show that

$$\text{(i)}\quad x^{n+1}\frac{d^{2n+1}(x^n y)}{dx^{2n+1}} + X^{n+1}\frac{d^{2n+1}(X^n y)}{dX^{2n+1}} = 0,$$

$$\text{(ii)}\quad x^{n+1}\frac{d^n(x^{n-1}y)}{dx^n} = (-)^n\frac{d^n y}{dX^n}.$$

8. Prove the equivalence of the pairs of operators

(i) $(x^2D)^n x^{1-n} = x^{n+2}D^{n+1}$,

(ii) $(x^{-1}D)^n x^{2n+1}(x^{-1}D)^{n+1} = D^{2n+1}$,

(iii) $(x^{\frac12}D)^{2n+1} = D^n x^{n+\frac12}D^{n+1}$,

(iv) $(x^{n-1/r}D^n)^r = x^{n-1}(x^{1-1/r}D)^{nr}$,

(v) $x^p\{x^{pq-1}(x^{1-p}D)^q\}^p = x^q\{x^{pq-1}(x^{1-q}D)^p\}^q$,

(vi) $\left\{x^{pq-p}\left(\dfrac{d}{dx_q}\right)^p\right\}^q = x^{pq-p}\left\{\left(\dfrac{d}{dx_p}\right)^q\right\}^p \qquad \left(x_p, x_q = \dfrac{x^p}{p}, \dfrac{x^q}{q}\right)$.

9. Prove the equivalence of the pairs of operators

$$\text{(i)}\quad \left(D+\frac{2n}{x}\right)\left(D+\frac{2n-2}{x}\right)\dots\left(D+\frac{2}{x}\right)D\left(D-\frac{2}{x}\right)\dots\left(D-\frac{2n-2}{x}\right)\left(D-\frac{2n}{x}\right) = D^{2n+1},$$

$$\text{(ii)}\quad \left(D+\frac{2n+1}{x}\right)\left(D+\frac{2n-1}{x}\right)\dots\left(D+\frac{1}{x}\right)\left(D-\frac{1}{x}\right)\dots\left(D-\frac{2n-1}{x}\right)\left(D-\frac{2n+1}{x}\right) = D^{2n+2},$$

$$\text{(iii)}\quad \left(D+\frac{1}{x}\right)^n\left(D-\frac{n}{x}\right) = D^{n+1}.$$

10. If $a, b, c, \dots$ are positive integers and we write

$$\Delta_a = x^a D^a x^a, \quad \text{etc.},$$

prove that $\Delta_a, \Delta_b, \Delta_c, \dots$ are commutative operators and that

$$\Delta_a\Delta_b\Delta_c\dots = \Delta_{a+b+c+\dots}.$$

If m, n are positive integers and $m > n$, prove the equivalence of operators

$$x^{m-n}D^m x^{m+n}D^n = D^n x^{m+n}D^m x^{m-n}.$$

11. Show that the two functions

$$x^{n+1}(x^{-1}D)^{n+1}e^{\pm x}$$

satisfy the differential equation

$$\frac{d^2y}{dx^2} - y = \frac{n(n+1)}{x^2}y,$$

and that the three functions

$$x^{3n+2}(x^{-2}D)^n x^{3m-1}(x^{-2}D)^m x^{-1}\{e^x, \quad e^{-\frac12 x}\cos\tfrac12 x\sqrt3, \quad e^{-\frac12 x}\sin\tfrac12 x\sqrt3\}$$

satisfy the differential equation

$$\frac{d^3y}{dx^3} - \frac{3(m+n)}{x}\frac{d^2y}{dx^2} + \frac{3m(3n+1)}{x^2}\frac{dy}{dx} - y = 0.$$

12. If f, g are polynomial functions of their arguments, show that

$$\lim_{x\to 0} f(D)\, g(x) = \lim_{x\to 0} g(D)\, f(x).$$

13. If P is a polynomial operator and $a_1, a_2, \ldots, a_n$ are distinct numbers, show that any solution of
$$(P-a_1)(P-a_2)\ldots(P-a_n)y = 0$$
can be written in the form

$$y = b_1\eta_1 + b_2\eta_2 + \ldots + b_n\eta_n,$$

where η_r is a solution of $(P-a_r)y = 0$.

14. If P is a polynomial operator of order n, and $\eta_1, \ldots, \eta_n$ are n linearly distinct solutions of $P^m y = 0$ not also solutions of $P^{m-1}y = 0$, show that a complete set of mn linearly distinct solutions of $P^m y = 0$ is given by

$$\begin{array}{c} \eta_1, \ \ldots, \ \eta_n, \\ P\eta_1, \ \ldots, \ P\eta_n, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ P^{m-1}\eta_1, \ \ldots, \ P^{m-1}\eta_n. \end{array}$$

15. If A is the alternant of P, Q, prove that the alternant of P, Q^n is

$$\sum_{r=0}^{n-1} Q^r A\, Q^{n-r-1}.$$

Prove that, if $P, f(Q)$ are commutative, so are P, Q.

16. (i) If f, g are polynomial functions of their arguments, show that

$$f\left(x+\frac{\partial}{\partial y}\right)g(y) = g\left(y+\frac{\partial}{\partial x}\right)f(x).$$

(ii) If α_r, β_s denote functions of x, and if

$$f(x,y) = \alpha_0 y^m + \alpha_1 y^{m-1} + \ldots + \alpha_m, \qquad g(x,y) = \beta_0 y^n + \beta_1 y^{n-1} + \ldots + \beta_n,$$

show that
$$\sum_{r=1}^{\infty} \frac{1}{r!}\left(\frac{\partial^r f}{\partial x^r}\frac{\partial^r g}{\partial y^r} - \frac{\partial^r g}{\partial x^r}\frac{\partial^r f}{\partial y^r}\right) = 0,$$

if and only if the operators

$$\alpha_0 D^m + \alpha_1 D^{m-1} + \ldots + \alpha_m, \qquad \beta_0 D^n + \beta_1 D^{n-1} + \ldots + \beta_n,$$

are commutative.

(iii) If the operators $\sum a_{rs} x^r D^s$, $\sum b_{rs} x^r D^s$ are commutative, where a_{rs}, b_{rs} are constants, show that the operators $\sum a_{rs} x^s D^r$, $\sum b_{rs} x^s D^r$ are also commutative.

17. If P, Q, R are mutually commutative, prove that, given p, we can find q, r such that the equations

$$(P-p)y = 0, \qquad (Q-q)y = 0, \qquad (R-r)y = 0$$

have a common solution.

18. Show that the operators

$$P_r = (xD^{r+1} - mD^r)x^{-1}D^m \qquad (r = 0, 1, \ldots, m)$$

form a set of $m+1$ mutually commutative operators and that the equations

$$(P_r - c^{m+r+1})y = 0 \qquad (r = 0, 1, \ldots, m)$$

have a common solution $\quad y = (cx-m)e^{cx}$.

19. If $P\,(\equiv D^m+\dots)$ is self-adjoint, prove that the solutions $\eta_1,\dots,\eta_m$ of the equation $Py=0$ obey relations

$$\sum_{r=1}^{m}\sum_{s=1}^{m} a_{rs}(D^p\eta_r)(D^q\eta_s) = 0 \qquad (p+q<m-1),$$
$$= \pm 1 \qquad (p+q=m-1). \quad \text{[Darboux.]}$$

If $P\,(\equiv D^m+\dots)$ and $Q\,(\equiv D^{2n}+\dots)$ are commutative and each self-adjoint and if η_r is now the common solution of

$$Py=0=(Q-q_r)y \qquad (r=1,\dots,m),$$

prove that the above relations simplify to

$$\sum a_r(D^p\eta_r)(D^q\eta_r) = 0,\ \pm 1.$$

20. Show that the expression Py becomes an exact derivative when multiplied by any function that is annihilated by the adjoint operator P'.

21. If $|a_{rs}|$ denotes the determinant whose typical constituent is a_{rs}, prove that

(i) in the field of n^2 variables (x_{rs})

$$|\partial/\partial x_{rs}|\,|x_{rs}| = n!,$$

(ii) in the field of n variables $(x_1,\dots,x_n)$

$$|\partial_r^{n-s}|\,|x_r^{n-s}| = n!!$$

Prove that in the field of n variables $(x_1,\dots,x_n)$ the result of operating on the circulant whose first row is $x_1,x_2,\dots,x_n$ by the circulant whose first row is $\partial_1,\partial_2,\dots,\partial_n$ is n^n.

22. If
$$\theta \equiv x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}+z\frac{\partial}{\partial z},$$

show that $\quad \theta(\theta-1)\dots(\theta-n+1) = \sum \dfrac{n!}{\alpha!\,\beta!\,\gamma!}x^\alpha y^\beta z^\gamma\left(\dfrac{\partial}{\partial x}\right)^\alpha\left(\dfrac{\partial}{\partial y}\right)^\beta\left(\dfrac{\partial}{\partial z}\right)^\gamma,$

where the summation is taken over all positive integer or zero solutions of the equation
$$\alpha+\beta+\gamma=n.$$

23. Prove that the operator

$$\frac{\partial(\xi,\eta)}{\partial(x_2,x_3)}\frac{\partial}{\partial x_1}+\frac{\partial(\xi,\eta)}{\partial(x_3,x_1)}\frac{\partial}{\partial x_2}+\frac{\partial(\xi,\eta)}{\partial(x_1,x_2)}\frac{\partial}{\partial x_3}$$

is self-adjoint, and show that any self-adjoint operator $\alpha_1\partial_1+\alpha_2\partial_2+\alpha_3\partial_3$ can be put into this form by proper choice of ξ,η.

24. If the p operators

$$\alpha_r+\alpha_{r1}\partial_1+\dots+\alpha_{rn}\partial_n \qquad (r=1,\dots,p)$$

are commutative, prove that by proper choice of variables $(\xi_1,\dots,\xi_n)$ they are reducible to
$$\frac{\partial\alpha}{\partial\xi_r}+\frac{\partial}{\partial\xi_r} \qquad (r=1,\dots,p),$$

where α is the same for every operator.

25. Prove that the pair of operators

$$\frac{\partial}{\partial\xi_r}, \qquad J^{-1}\frac{\partial J}{\partial\xi_r}-\frac{\partial}{\partial\xi_r}, \quad \text{where } J=\frac{\partial(\xi_1,\dots,\xi_n)}{\partial(x_1,\dots,x_n)},$$

are adjoint, when transformed from variables $(\xi_1,\dots,\xi_n)$ to variables $(x_1,\dots,x_n)$.

26. If ξ is homogeneous of degree m in $x_1, \ldots, x_n$ and if $P = \sum x_r \partial_r$, prove that

$$P(P-m)\ldots(P-mr+m)f(\xi) = m^r \xi^r f^{(r)}(\xi).$$

If $Q \equiv \sum \alpha_r \partial_r$ where every α_r is homogeneous of the same degree m, prove that the alternant of P, Q is $(m-1)Q$, and of P, Q^s is $s(m-1)Q^s$. Prove also that

$$f(P)Q = Qf(P+m-1),$$
$$f(Q)P = Pf(Q)-(m-1)Qf'(Q).$$

27. Show that, if p, q are integers, the series of successive alternants of

$$\sum x_r^p \partial_r, \qquad \sum x_r^q \partial_r$$

is endless, unless one of $p-1$, $q-1$, $p+q-2$ is zero. In these cases reduce them to the special forms of (63).

28. Show that the operators

$$x_1 \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_2} + \ldots + x_n \frac{\partial}{\partial x_n},$$
$$x_2 \frac{\partial}{\partial x_1} + x_3 \frac{\partial}{\partial x_2} + \ldots + x_1 \frac{\partial}{\partial x_n},$$
$$x_3 \frac{\partial}{\partial x_1} + x_4 \frac{\partial}{\partial x_2} + \ldots + x_2 \frac{\partial}{\partial x_n}$$

are commutative and reduce them to the special forms

$$\frac{\partial}{\partial \xi_1}, \frac{\partial}{\partial \xi_2}, \frac{\partial}{\partial \xi_3}.$$

Hence or otherwise obtain the complete set of n commutative operators of which the three given operators form part.

29. Obtain the alternant E of the operators

$$\Omega = x_0 \frac{\partial}{\partial x_1} + 2x_1 \frac{\partial}{\partial x_2} + \ldots + px_{p-1} \frac{\partial}{\partial x_p},$$
$$O = x_p \frac{\partial}{\partial x_{p-1}} + 2x_{p-1} \frac{\partial}{\partial x_{p-2}} + \ldots + px_1 \frac{\partial}{\partial x_0},$$

and show that the three operators Ω, O, E have no alternant linearly distinct from themselves.

Show that the operators

$$E, \qquad x_p^{1+2/p}(kx_{p-1}E + x_p\Omega), \qquad x_0^{1+2/p}(k'x_1 E - x_0 O)$$

are commutative, if $k+k'=1$.

30. If m, n are positive integers and P denotes the operator $x\partial_y\, y\partial_x$, show that

$$\left\{1 - \frac{P}{(1!)^2} + \frac{P^2}{(2!)^2} - \frac{P^3}{(3!)^2} + \ldots \text{ to } \infty\right\} x^m y^n = \begin{cases} 0 & (m > n), \\ (-)^m \dfrac{n!\,x^m y^n}{m!\,(n-m)!} & (m < n). \end{cases}$$

31. If $f(x, y)$ is a polynomial, show that

$$\exp\left(ax\frac{\partial}{\partial y}\right)\exp\left(by\frac{\partial}{\partial x}\right)f(x,y) = f\{(1+ab)x+by,\ ax+y\},$$

and effect on $f(x, y)$ the operation

$$\exp\left(-ax\frac{\partial}{\partial y}\right)\exp\left(-by\frac{\partial}{\partial x}\right)\exp\left(ax\frac{\partial}{\partial y}\right)\exp\left(by\frac{\partial}{\partial x}\right).$$

32. If $ps-qr = 1$, show that the operator

$$\exp\left(\frac{(s-1)y}{r}\frac{\partial}{\partial x}\right)\exp\left(rx\frac{\partial}{\partial y}\right)\exp\left(\frac{(p-1)y}{r}\frac{\partial}{\partial x}\right)$$

on any polynomial replaces x, y by $px+qy$, $rx+sy$.

Obtain an operator of this type that will change a polynomial $f(x, y)$ into $f(y, -x)$.

33. Show that, applied to any polynomial, the operators

$$\exp\left(\lambda\mu x\frac{\partial}{\partial y}\right)\exp\left(\frac{(\lambda-1)y}{\lambda\mu}\frac{\partial}{\partial x}\right)\exp\left(-\mu x\frac{\partial}{\partial y}\right)\exp\left(\frac{(1-\lambda)y}{\mu}\frac{\partial}{\partial x}\right),$$

$$\sum_{r=0}^{\infty}\frac{(\lambda-1)^r x^r}{r!}\frac{\partial^r}{\partial x^r}\sum_{r=0}^{\infty}\frac{(\lambda^{-1}-1)^r y^r}{r!}\frac{\partial^r}{\partial y^r}$$

are equivalent, and that their effect is to replace x, y by λx, $\lambda^{-1}y$.

34. Prove that, if f is a polynomial,

$$z = \exp\left(x\frac{\partial^n}{\partial y^n}\right)f(y)$$

is a solution of the differential equation

$$\frac{\partial z}{\partial x} = \frac{\partial^n z}{\partial y^n},$$

and so are the n expressions

$$z_p = \sum_{r=0}^{\infty}\frac{y^{nr+p}}{(nr+p)!}\frac{\partial^r}{\partial x^r}f(x) \qquad (p = 0,\ldots, n-1).$$

35. If $f(t)$, $g(t)$ are polynomials in t, show that the partial differential equation

$$\frac{\partial z}{\partial x} = f\left(\frac{\partial}{\partial y}\right)z$$

has the solution
$$z = \exp\left\{xf\left(\frac{\partial}{\partial y}\right)\right\}g(y).$$

36. Show that the partial differential equation

$$\frac{\partial^2 z}{\partial x\partial y} = u(x, y)z$$

has the *symbolical* solution

$$\sum_{r=0}^{\infty}\left(\frac{1}{\partial_x\partial_y}u\right)^r\{f(x)+g(y)\},$$

where f, g are arbitrary functions.

37. If u, v, w denote functions of x, y, show that the partial differential equation

$$\frac{\partial^2 z}{\partial x\partial y}+u\frac{\partial z}{\partial x}+v\frac{\partial z}{\partial y}+w = 0$$

has a *symbolical* solution

$$z = e^{\alpha}\sum_{r=0}^{\infty}\left\{\frac{1}{\partial_x\partial_y}(uv+u_x-w)-\frac{1}{\partial_y}(u+\alpha_y)\right\}^r f(x),$$

where $\alpha_x = -v$ and $f(x)$ is an arbitrary function of x.

38. What conditions, if any, must be imposed on m, n, p in order that, in the notation of § 10,

$$y_m y_n, \qquad y_m y_{n+p} + y_n y_{m+p}, \qquad y_m y_{n+p} + y_n y_{p+m} + y_p y_{m+n}$$

may be exact derivatives. When the conditions are satisfied, obtain the corresponding primitives. Show also that one of

$$y_m y_{n+p} \pm y_n y_{m+p}$$

is always an exact derivative, and that, if $m+n$ and p are both odd,

$$y_m y_{n+p} + y_n y_{m+p}$$

is an exact second derivative.

39. Show that the fractional expression

$$y y_2 / y_1^2$$

is annihilated by the operator Δ of (68). Is it an exact derivative? At what point may the argument of § 10 fail in being applied to fractional expressions?

40. If $$x^r F(x, y, y_1, \dots) \qquad (r = 0, 1, \dots, n-1)$$

are all exact derivatives, show that F itself is an exact nth derivative, and conversely.

If $$\Delta_r = \partial_r - (r+1)D\partial_{r+1} + \frac{(r+1)(r+2)}{2!}D^2\partial_{r+2} - \dots,$$

show that an exact nth derivative is annihilated by $\Delta, \Delta_1, \dots, \Delta_{n-1}$, and, conversely, that a polynomial $F(x, y, y_1, \dots)$ annihilated by each of these n operators is an exact nth derivative.

41. If $F(x, y, y_1, \dots)$ is an exact derivative, show that $\partial F/\partial x$, $\partial F/\partial y$ are also exact derivatives.

If F is an exact nth derivative, show that $\partial F/\partial y_r$ $(r < n)$ is an exact $(n-r)$th derivative.

Obtain in each case an expression for the corresponding primitive.

42. If there are two *dependent* variables y, z, prove that the corresponding annihilator of exact derivatives $F(x, y, z, y_1, z_1, \dots)$ is

$$\Delta_{yz} = \frac{\partial}{\partial y} + \frac{\partial}{\partial z} - D\left(\frac{\partial}{\partial y_1} + \frac{\partial}{\partial z_1}\right) + D^2\left(\frac{\partial}{\partial y_2} + \frac{\partial}{\partial z_2}\right) - \dots .$$

If P, Q are two ordinary differential operators and P', Q' their adjoints, prove that $$\Delta_{yz}(yPz - zQy) = (P - Q')z + (P' - Q)y.$$

XIV

EXPANSION IN POWER-SERIES

1. The practical problem

In chapter VIII we considered, at some length, the question of the expansibility of functions in power-series. The approach to the question was essentially theoretical: we were concerned with existence theorems which should decide what classes of functions are expansible in power-series and under what conditions. But those theorems, once they have decided the expansibility of a particular function, give little help in the practical problem of obtaining its expansion in concise and explicit terms.

It is that practical problem which we are now to consider, completing the earlier theory with convenient methods of expanding specific functions. In applying these methods we shall not need to make a preliminary test for expansibility. For we have, in any case, to settle whether a particular method of expansion is applicable: if it is, then the expansion itself answers the question of expansibility.

The problem, of course, is not essentially a problem of the Differential Calculus: algebraic methods are used with much effect in certain cases, but their general scope is limited, and the Differential Calculus, as a rule, proves a much more powerful instrument.

2. Taylor's series and Lagrange's series

Now, by chapter VIII (11), we know at the outset that the expansion, when it exists, is necessarily a Taylor's series of the function. The expansion can therefore be written down when we can evaluate the Taylor's series, that is to say, when we have a convenient formula for the nth derivative of the function. As we saw in chapter V § 2, such cases are very rare. We may cite e^x, $\sin x$, or more generally $e^{ax}\sin(bx+c)$; and certain algebraic expressions that can be expanded equally well by elementary algebraic methods. Taking e^x as sufficiently illustrative, the argument would run thus:

$$|D^n e^x/n!|^{1/n} = |e^x/n!|^{1/n} < e^{x/n} < e, \quad \text{if } n > x.$$

Thus $$|D^n e^x/n!|^{1/n}$$

is bounded in any finite interval and therefore, by chapter VIII (19), e^x is equal to any Taylor series,† and, in particular, to the Maclaurin-

† The expansibility of e^x has already been established (chapter VIII § 9). It is repeated here for completeness in an illustrative argument.

Taylor series in any finite interval of convergence. Since

$$[D^n e^x]_{x=0} = 1,$$

we at once have the expansion

$$e^x = \sum_0^\infty \frac{x^n}{n!}. \tag{1}$$

Lagrange's series arises as a by-product of Taylor's series, applicable to functions defined implicitly. Here, too, the actual expansion itself forms an integral part of the existence theory, and the practical problem is solved whenever certain nth derivatives have a convenient form. The theory is covered by theorem (10) of chapter XI, but is stated explicitly in example 20 of chapter VIII. It is briefly as follows.

If y is defined as a function of x by the equation

$$y = a + x\phi(y),$$

where, when y is many-valued, that branch is taken which converges to a at $x = 0$, then we can expand $f(y)$ in the series

$$f(y) = f(a) + \sum_{n=1}^\infty \frac{x^n}{n!}\frac{d^{n-1}}{da^{n-1}}[\{\phi(a)\}^n f'(a)], \tag{2}$$

provided that $f(y)$, $\phi(y)$ are analytic functions of y, and $1 - x\phi'(y)$ does not vanish.

Here again we do not get a convenient form for the typical coefficient in the expansion except in a few special cases of which the following are sufficiently representative.

Take $\phi(y) = y^m$, $f(y) = y^p$. There is no loss of generality if we put $a = 1$, so that y is a root of the equation

$$xy^m - y + 1 = 0, \tag{3}$$

and the series is

$$y^p = 1 + p\sum_1^\infty \frac{(mn+p-1)(mn+p-2)\ldots(mn+p-n+1)}{n!}x^n. \tag{4}$$

If now we take $f(y) = \log y$, which of course corresponds to the limiting case $p = 0$, we get

$$\log y = \sum_1^\infty \frac{(mn-1)(mn-2)\ldots(mn-n+1)}{n!}x^n. \tag{5}$$

We should perhaps note that $p = 1$ in (4) gives y itself as

$$y = 1 + m\sum_1^\infty \frac{(mn-1)(mn-2)\ldots(mn-n+2)}{(n-1)!}x^n. \tag{6}$$

If m is a positive integer, we thus obtain one root of the trinomial m-ic (3). We could set about to apply this method to determine the remaining roots, but the complete solution of (3) is more easily discussed when we can call in the methods of a subsequent section.

Again, take $\phi(y) = e^y$, $f(y) = e^{cy}$, $a = 0$, so that y is a root of the equation

$$y = xe^y,$$

and the series is
$$e^{cy} = 1 + c\sum_1^\infty \frac{(n+c)^{n-1}x^n}{n!}. \tag{7}$$

We may eliminate x and write this explicitly as

$$e^{cy} = 1 + c\sum_1^\infty \frac{(n+c)^{n-1}y^n e^{-ny}}{n!}. \tag{8}$$

Had we taken $f(y) = y$, we should have obtained the similar expansion

$$y = \sum_1^\infty \frac{n^{n-2}x^n}{(n-1)!}, \tag{9}$$

which is therefore the expansion of $y(x)$, the function inverse to

$$x(y) = ye^{-y}.$$

If, in (7), we write $e^y = y'$, and drop accents, we have

$$y^c = 1 + c\sum_1^\infty \frac{(n+c)^{n-1}x^n}{n!}, \tag{10}$$

where $x = y^{-1}\log y$: let us say

$$y^c = 1 + c\sum_1^\infty \frac{(n+c)^{n-1}(y^{-1}\log y)^n}{n!}. \tag{11}$$

In particular, $c = 1$ gives

$$y = 1 + \sum_1^\infty \frac{(n+1)^{n-1}x^n}{n!}, \tag{12}$$

which is the expansion of the function $y(x)$ inverse to the function

$$x(y) = y^{-1}\log y.$$

It is obvious that the series (9), (12) are closely connected.

3. The integration of a known series

From the point of view of this book the method of this section is presented rather as the differentiation of an unknown series, but the distinction is, of course, immaterial. We compare a series

$$s(x) = \sum_{n=0}^\infty a_n x^n$$

with the series $$S(x) \equiv \sum_{n=1}^{\infty} na_n x^{n-1}$$

obtained from differentiating $s(x)$ term by term. The known facts are these.†

If $S(x)$ converges at $x = \xi$, then $s(x)$ also converges at $x = \xi$. Moreover, both series converge uniformly over the closed interval $(0, \xi)$ and, throughout this interval, $S(x)$ is the derivative of $s(x)$. It is sufficient, therefore, to write merely

$$S(\xi) = s'(\xi).$$

The intervals of convergence of the two series are the same, with the possible exception of one or both end-points: here $s(x)$ may converge although $S(x)$ do not. At such an end-point $s(x)$ must be examined independently. But, by what has been said, $s(x)$ converges uniformly up to this end-point and so its sum is continuous there. If, as is generally the case, the expression we have found for $s(x)$ as the integral of $S(x)$ is also continuous at this end-point, then the expansion still holds at the end-point.

The constant term a_0 in $s(x)$ is absent from $S(x)$: it is, in fact, the arbitrary constant of integration. We determine it by setting $x = 0$ The following examples show this method applied to the expansion of certain elementary inverse functions.

Write $$s(x) = \log(1+x).$$

Then $$s'(x) = (1+x)^{-1} = 1-x+x^2 - \dots \quad \text{to } \infty,$$

and $$s(x) = a_0+x-\tfrac{1}{2}x^2+\tfrac{1}{3}x^3- \dots .$$

But $a_0 = s(0) = 0$, and so

$$\log(1+x) = x-\tfrac{1}{2}x^2+\tfrac{1}{3}x^3- \dots \quad \text{to } \infty. \tag{13}$$

The differentiated series converges in the open interval $)-1, 1($. The series (13) converges at $x = 1$, though not at $x = -1$. But $\log(1+x)$ is continuous at $x = 1$, and (13) therefore holds throughout the half-open interval $)-1, 1)$.

Write $$s(x) = \tan^{-1} x.$$

Then $$s'(x) = (1+x^2)^{-1} = 1-x^2+x^4- \dots \quad \text{to } \infty.$$

and $$s(x) = a_0+x-\tfrac{1}{3}x^3+\tfrac{1}{5}x^5- \dots .$$

If $\tan^{-1} x$ denotes, as usual, the principal value of the inverse function, then $a_0 = 0$ and we have

$$\tan^{-1}x = x-\tfrac{1}{3}x^3+\tfrac{1}{5}x^5- \dots \quad \text{to } \infty. \tag{14}$$

† Cf. chapter IV (74).

Here again the differentiated series converges only in $)-1,1($, but $s(x)$ itself converges at $x=1$, though not at $x=-1$. Now $\tan^{-1}x$ is continuous at $x=1$, and therefore (14) holds throughout $)-1,1)$.

The corresponding formula for the inverse hyperbolic tangent

$$\tanh^{-1}x = \log\frac{1+x}{1-x} = x+\tfrac{1}{3}x^3+\tfrac{1}{5}x^5+\ldots \tag{15}$$

is at once deducible from (13). It is valid throughout $(-1,1($.

Write $$s(x) = \sin^{-1}x.$$

Then $$s'(x) = (1-x^2)^{-\frac{1}{2}} = 1+\frac{1}{2}x^2+\frac{1.3}{2.4}x^4+\ldots \quad \text{to } \infty,$$

and $$s(x) = a_0+x+\frac{1}{2}.\frac{x^3}{3}+\frac{1.3}{2.4}.\frac{x^5}{5}+\ldots.$$

If $\sin^{-1}x$ denotes, as usual, the principal value of the inverse function, then $a_0=0$ and we have

$$\sin^{-1}x = x+\frac{1}{2}.\frac{x^3}{3}+\frac{1.3}{2.4}.\frac{x^5}{5}+\ldots \quad \text{to } \infty. \tag{16}$$

Again, the differentiated series converges in the open interval $)-1,1($. Now, if u_n denote the coefficient of x^{2n-1} in (16), we have

$$\frac{u_{n+1}}{u_n} = \frac{(2n-1)^2}{2n(2n+1)} = 1-\tfrac{3}{2}n^{-1}+On^{-2},$$

and therefore, by a standard test,† $\sum u_n$ converges. Thus (16) converges at $x=\pm 1$, and $\sin^{-1}x$ is continuous at these two points. Hence (16) holds throughout the closed interval $(-1,1)$.

Similarly for the inverse hyperbolic sine we shall have

$$\sinh^{-1}x = \log\{x+\sqrt{(x^2+1)}\} = x-\frac{1}{2}.\frac{x^3}{3}+\frac{1.3}{2.4}.\frac{x^5}{5}-\ldots \quad \text{to } \infty, \tag{16'}$$

which, again, is valid if $|x|\leqslant 1$.

We cannot, of course, expand

$$\cosh^{-1}x = \log\{x+\sqrt{(x^2-1)}\}$$

in ascending powers of x, since this is real only if $x>1$. We consider instead

$$\operatorname{sech}^{-1}x = \log\{1+\sqrt{(1-x^2)}\}-\log x.$$

Take $$s(x) = \log\{1+\sqrt{(1-x^2)}\}.$$

Then $$s'(x) = \frac{-x/\sqrt{(1-x^2)}}{1+\sqrt{(1-x^2)}} = -\frac{1-\sqrt{(1-x^2)}}{x\sqrt{(1-x^2)}} = \frac{1-(1-x^2)^{-\frac{1}{2}}}{x}.$$

Thus $$s'(x) = -\frac{1}{2}x-\frac{1.3}{2.4}x^3-\frac{1.3.5}{2.4.6}x^5-\ldots \quad \text{to } \infty.$$

† Cf. Bromwich, *Infinite Series* (1926), 39–40.

Now $s(0) = \log 2$, and so

$$\log\{1+\sqrt{(1-x^2)}\} = \log 2 - \frac{1}{2}.\frac{x^2}{2} - \frac{1.3}{2.4}.\frac{x^4}{4} - \frac{1.3.5}{2.4.6}.\frac{x^6}{6} - \dots \quad \text{to } \infty. \quad (17)$$

As with $\sin^{-1}x$, we can show that this is valid, if $|x| \leqslant 1$. Thus

$$\begin{aligned}\operatorname{sech}^{-1}x &= \log\{1+\sqrt{(1-x^2)}\} - \log x \\ &= \log\left(\frac{2}{x}\right) - \frac{1}{2}.\frac{x^2}{2} - \frac{1.3}{2.4}.\frac{x^4}{4} - \frac{1.3.5}{2.4.6}.\frac{x^6}{6} - \dots \quad \text{to } \infty, \qquad (18)\end{aligned}$$

which is valid only if $1 > x > 0$. Similarly

$$\begin{aligned}\operatorname{cosech}^{-1}x &= \log\{1+\sqrt{(1+x^2)}\} - \log x \\ &= \log\left(\frac{2}{x}\right) + \frac{1}{2}.\frac{x^2}{2} - \frac{1.3}{2.4}.\frac{x^4}{4} + \frac{1.3.5}{2.4.6}.\frac{x^6}{6} - \dots \quad \text{to } \infty. \qquad (19)\end{aligned}$$

From (17) we can deduce immediately the expansions of

$$\log\{1-\sqrt{(1-x^2)}\}, \qquad \log\{\sqrt{(1+x)} \pm \sqrt{(1-x)}\}.$$

4. Two important trigonometric series

We may consider here two other important trigonometrical series of a different type. For the first, write

$$f(x) \equiv x\sin\alpha + \tfrac{1}{2}x^2\sin 2\alpha + \tfrac{1}{3}x^3\sin 3\alpha + \dots \quad \text{to } \infty.$$

By Abel's principle† this is known to converge when $x = 1$; it therefore converges uniformly over the closed interval $(0, 1)$, and, in particular, is continuous as $x \to 1-$.

Term-by-term differentiation gives

$$f'(x) = \sin\alpha + x\sin 2\alpha + x^2\sin 3\alpha + \dots,$$

which converges in the half-open interval $(0, 1($. There is therefore uniform convergence over any closed interval $(0, 1-\delta)$ where $0 < \delta < 1$, and the term-by-term differentiation is legitimate throughout every such interval, i.e. at all points of $(0, 1($.

Elementary methods give for the actual sum

$$f'(x) = \frac{\sin\alpha}{1-2x\cos\alpha+x^2} = \frac{d}{dx}\left\{\tan^{-1}\left(\frac{x-\cos\alpha}{\sin\alpha}\right)\right\},$$

where, to avoid ambiguity, we define the inverse tangent to mean the appropriate angle lying between $\pm\frac{1}{2}\pi$. Thus

$$f(x) = \tan^{-1}\left(\frac{x-\cos\alpha}{\sin\alpha}\right) - C \qquad (0 \leqslant x < 1)$$

where C is a constant. Put $x = 0$, and we get

$$\begin{aligned}C &= \tan^{-1}(-\cot\alpha) \\ &= \alpha + (n-\tfrac{1}{2})\pi, \quad \text{where } n \text{ is some integer or zero.}\end{aligned}$$

† Cf. Bromwich, *Infinite Series* (1926), 57–60 (§§ 20, 22).

If we now restrict α to the open interval $)0, \pi($, we have $n = 0$, and so

$$f(x) = \tan^{-1}\left(\frac{x-\cos\alpha}{\sin\alpha}\right)-\alpha+\tfrac{1}{2}\pi \qquad \begin{pmatrix} 0 \leqslant x < 1 \\ 0 < \alpha < \pi \end{pmatrix}.$$

Now we have seen that $f(x)$ is continuous as $x \to 1-$, and the inverse tangent is also continuous there, since $\sin\alpha \neq 0$. We may therefore proceed to the limit and get

$$f(1) = \tan^{-1}(\tan\tfrac{1}{2}\alpha)-\alpha+\tfrac{1}{2}\pi = \tfrac{1}{2}(\pi-\alpha),$$

taking $\tan^{-1}(\tan\frac{1}{2}\alpha) = \frac{1}{2}\alpha$, since $0 < \frac{1}{2}\alpha < \frac{1}{2}\pi$. Thus

$$\tfrac{1}{2}(\pi-\alpha) = \sin\alpha+\tfrac{1}{2}\sin 2\alpha+\tfrac{1}{3}\sin 3\alpha+\ldots \quad \text{to } \infty \quad (0 < \alpha \leqslant \pi),$$

since we may include the trivial case $\alpha = \pi$.

If we write $2\pi-\alpha$ for α, both sides of this expression change sign; the result therefore holds over the wider interval $0 < \alpha < 2\pi$. If we write $2n\pi+\alpha$ for α, where n is an integer, the series is unaltered, but the expression on the left is diminished by $n\pi$. We therefore have, for any α,

$$\tfrac{1}{2}(\pi-\alpha)+n\pi = \sin\alpha+\tfrac{1}{2}\sin 2\alpha+\tfrac{1}{3}\sin 3\alpha+\ldots \quad \text{to } \infty, \tag{20}$$

where $2n\pi < \alpha < (2n+2)\pi$.

When $\alpha = 2n\pi$, the series still converges but to the sum zero. Thus, as we have already explained in an earlier chapter,† the sum-function of the series is discontinuous at all the points $\alpha = 2n\pi$, where n is an integer or zero.

We can deal similarly with the allied series

$$g(x) = x\cos\alpha+\tfrac{1}{2}x^2\cos 2\alpha+\tfrac{1}{3}x^3\cos 3\alpha+\ldots \quad \text{to } \infty.$$

This again is uniformly convergent in $(0,1)$ and therefore continuous as $x \to 1-$. The differentiated series

$$g'(x) = \cos\alpha+x\cos 2\alpha+x^2\cos 3\alpha+\ldots$$

converges in $(0, 1($ to

$$\frac{\cos\alpha-x}{1-2x\cos\alpha+x^2} = -\frac{1}{2}\frac{d}{dx}\log(1-2x\cos\alpha+x^2),$$

the differentiation being valid throughout this half-open interval. We therefore have

$$g(x) = C-\tfrac{1}{2}\log(1-2x\cos\alpha+x^2) \qquad (0 \leqslant x < 1)$$

where C is a constant, which, by putting $x = 0$, is found to vanish. As we have seen, $g(x)$ is continuous as $x \to 1-$, and the logarithm is also continuous there, if $\alpha \neq 2n\pi$. We may therefore proceed to this limit, which gives us

$$\log\tfrac{1}{2}\operatorname{cosec}\tfrac{1}{2}\alpha = \cos\alpha+\tfrac{1}{2}\cos 2\alpha+\tfrac{1}{3}\cos 3\alpha+\ldots \quad \text{to } \infty \quad (\alpha \neq 2n\pi). \tag{20'}$$

† Chapter II § 7, p. 36.

5. Expansion of certain products and powers

This method may also be applied in repetition to obtain the expansion of certain powers and products of inverse functions.† Thus write

$$S(x) \equiv x-(1+\tfrac{1}{2})x^2+(1+\tfrac{1}{2}+\tfrac{1}{3})x^3-\dots \quad \text{to } \infty.$$

The typical coefficient
$$1+\frac{1}{2}+\frac{1}{3}+\dots+\frac{1}{n}$$
diverges like $\log n$. The series $S(x)$ therefore converges in the open interval $|x|<1$. In that interval we have

$$S+xS = x-\tfrac{1}{2}x^2+\tfrac{1}{3}x^3-\dots = \log(1+x).$$

Thus
$$S(x) = \frac{\log(1+x)}{1+x} = \frac{d}{dx}\tfrac{1}{2}\{\log(1+x)\}^2,$$
and so
$$\tfrac{1}{2}\{\log(1+x)\}^2 = \frac{x^2}{2}-(1+\tfrac{1}{2})\frac{x^3}{3}+(1+\tfrac{1}{2}+\tfrac{1}{3})\frac{x^4}{4}-\dots \quad \text{to } \infty. \tag{21}$$

To consider convergence at $x = 1$ write

$$u_n \equiv \left(1+\frac{1}{2}+\dots+\frac{1}{n}\right)\frac{1}{n+1}.$$

Then
$$u_n-u_{n+1} = \left(1+\frac{1}{2}+\dots+\frac{1}{n}\right)\left(\frac{1}{n+1}-\frac{1}{n+2}\right)-\frac{1}{(n+1)(n+2)}$$
$$= \left(\frac{1}{2}+\dots+\frac{1}{n}\right)\frac{1}{(n+1)(n+2)} > 0.$$

Thus u_n tends monotonically to zero, and the alternating series $\sum (-)^n u_n$ is convergent. Hence (21) holds in the half-open interval $)-1, 1)$.

More generally, if $c_{n,r}$ denote the sum of products, r at a time, of the first n reciprocals $1, \frac{1}{2},\dots, n^{-1}$, we can prove that

$$\frac{\{\log(1+x)\}^{r+1}}{(r+1)!} = \frac{c_{r,r}}{r+1}x^{r+1}-\frac{c_{r+1,r}}{r+2}x^{r+2}+\frac{c_{r+2,r}}{r+3}x^{r+3}-\dots \quad \text{to } \infty. \tag{21$'$}$$

For, by definition of $c_{n,r}$, we have at once that

$$c_{r,r} = \frac{c_{r-1,r-1}}{r}, \qquad c_{n+1,r}-c_{n,r} = \frac{c_{n,r-1}}{n+1} \quad (n \geq r). \tag{22}$$

Differentiating both sides of (21′) we need to show that

$$\frac{\{\log(1+x)\}^r}{r!} = (1+x)\{c_{r,r}x^r-c_{r+1,r}x^{r+1}+c_{r+2,r}x^{r+2}-\dots\}$$
$$= c_{r,r}x^r-(c_{r+1,r}-c_{r,r})x^{r+1}+(c_{r+2,r}-c_{r+1,r})x^{r+2}-\dots$$
$$= \frac{c_{r-1,r-1}}{r}x^r-\frac{c_{r,r-1}}{r+1}x^{r+1}+\frac{c_{r+1,r-1}}{r+2}x^{r+2}-\dots, \qquad \text{by (22).}$$

† Direct algebraic methods of obtaining the same results will also suggest themselves.

This is again (21′) with $r-1$ written for r. The formula therefore follows by induction, since we have already proved it for $r = 2$ (and $r = 1$).

It remains to consider the convergence of the series on the right of (21′). We first prove inductively that

$$c_{n,r} < (1+\log n)^r/r! \tag{23}$$

This is true for $r = 1$, since, by a fundamental inequality,

$$\frac{1}{m+1} < \log\left(1+\frac{1}{m}\right), \tag{24}$$

and therefore $$1+\sum_{m=1}^{n-1}\frac{1}{m+1} < 1+\sum_{m=1}^{n-1}\log\left(1+\frac{1}{m}\right),$$

i.e. $$c_{n,1} < 1+\log n.$$

Again, supposing (23) true for $r-1$, we have from (22)

$$c_{n+1,r}-c_{n,r} < \frac{(1+\log n)^{r-1}}{(r-1)!(n+1)}$$

$$< \frac{r\log(1+1/n)(1+\log n)^{r-1}}{r!}, \quad \text{by (24)},$$

$$< \frac{\{1+\log(1+1/n)+\log n\}^r-(1+\log n)^r}{r!}.$$

Thus $$\{1+\log(n+1)\}^r/r!-c_{n+1,r} > (1+\log n)^r/r!-c_{n,r}.$$

Using this inductively we at length get that

$$\{1+\log(n+1)\}^r/r!-c_{n+1,r} > (1+\log r)^r/r!-c_{r,r}$$

$$> 0, \quad \text{since } c_{r,r} = 1/r!$$

Thus (23) is proved.

Hence, if $$u_n \equiv c_{r+n-1,r}/(n+r)$$

is the typical coefficient in (21′), we see from (23) that $u_n \to 0$ as $n \to \infty$. But $\sum u_n$ diverges, since evidently $c_{n,r} > 1/r!$, i.e. the series in (21′) diverges at $x = -1$ (as we should expect). At $x = 1$ this series becomes $$\sum (-)^{n-1}u_n.$$

Now

$$u_{n+1}-u_{n+2} = \frac{c_{r+n,r}}{r+n+1}-\frac{c_{r+n+1,r}}{r+n+2}$$

$$= c_{r+n,r}\left(\frac{1}{r+n+1}-\frac{1}{r+n+2}\right)-\frac{c_{r+n,r-1}}{(r+n+1)(r+n+2)}, \quad \text{by (22)},$$

$$= \frac{c_{r+n,r}-c_{r+n,r-1}}{(r+n+1)(r+n+2)}.$$

Hence
$$|u_{n+1}-u_{n+2}| < \frac{c_{r+n,r}+c_{r+n,r-1}}{(r+n+1)(r+n+2)},$$
and so, by (23), as $n \to \infty$,
$$|u_{n+1}-u_{n+2}| = \epsilon n^{-(1+a)} \qquad (1 > a > 0).$$
Thus the bracketed series $\sum (u_{2n+1}-u_{2n+2})$ converges, and therefore, since $u_n \to 0$, the full series $\sum (-)^n u_n$ also converges, and (21′) holds in the half-open interval $)-1, 1)$.

In like fashion we can show that
$$\tfrac{1}{2}(\tan^{-1}x)^2 = \frac{x^2}{2}-(1+\tfrac{1}{3})\frac{x^4}{4}+(1+\tfrac{1}{3}+\tfrac{1}{5})\frac{x^6}{6}-\dots \quad \text{to } \infty, \tag{25}$$
$$\tfrac{1}{2}\log(1+x^2)\tan^{-1}x = (1+\tfrac{1}{2})\frac{x^3}{3}-(1+\tfrac{1}{2}+\tfrac{1}{3}+\tfrac{1}{4})\frac{x^5}{5}+$$
$$+(1+\tfrac{1}{2}+\tfrac{1}{3}+\tfrac{1}{4}+\tfrac{1}{5}+\tfrac{1}{6})\frac{x^7}{7}-\dots \quad \text{to } \infty, \tag{26}$$
the expansions being valid over the half-open interval $)-1, 1)$. For further products or powers of $\tan^{-1}x$ and $\log(1+x^2)$ the expansions are less manageable.

A somewhat similar expansion is that of
$$\log(1+x)\log(1-x)-\log(1-x^2)$$
We have
$$\frac{d}{dx}\{\log(1+x)\log(1-x)\} = \frac{\log(1-x)}{1+x}-\frac{\log(1+x)}{1-x}.$$
Now, as above,
$$\frac{\log(1-x)}{1+x} = -x+(1-\tfrac{1}{2})x^2-(1-\tfrac{1}{2}+\tfrac{1}{3})x^3+\dots \quad \text{to } \infty$$
and so
$$\frac{\log(1-x)}{1+x}-\frac{\log(1+x)}{1-x} = -2\{x+(1-\tfrac{1}{2}+\tfrac{1}{3})x^3+(1-\tfrac{1}{2}+\tfrac{1}{3}-\tfrac{1}{4}+\tfrac{1}{5})x^5+\dots\}.$$
Thus
$$\log(1+x)\log(1-x)$$
$$= -\{x^2+(1-\tfrac{1}{2}+\tfrac{1}{3})\frac{x^4}{2}+(1-\tfrac{1}{2}+\tfrac{1}{3}-\tfrac{1}{4}+\tfrac{1}{5})\frac{x^6}{3}+\dots \quad \text{to } \infty\}, \tag{27}$$
and so
$$\log(1+x)\log(1-x)-\log(1-x^2)$$
$$= (\tfrac{1}{2}-\tfrac{1}{3})\frac{x^4}{2}+(\tfrac{1}{2}-\tfrac{1}{3}+\tfrac{1}{4}-\tfrac{1}{5})\frac{x^6}{3}+(\tfrac{1}{2}-\tfrac{1}{3}+\tfrac{1}{4}-\tfrac{1}{5}+\tfrac{1}{6}-\tfrac{1}{7})\frac{x^8}{4}+\dots \quad \text{to } \infty. \tag{28}$$

As an example of somewhat different type we can show that

$$\left(1+x+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!}\right)e^{-x}$$

$$= 1-\frac{x^{n+1}}{(n+1)!}+\frac{n+1}{1!}\frac{x^{n+2}}{(n+2)!}-\frac{(n+1)(n+2)}{2!}\frac{x^{n+3}}{(n+3)!}+\ldots \quad \text{to } \infty. \tag{29}$$

For, differentiating both sides, we have to show that

$$-\frac{x^n e^{-x}}{n!} = -\frac{x^n}{n!}+\frac{x^{n+1}}{1!\,n!}-\frac{x^{n+2}}{2!\,n!}+\ldots,$$

which is right. We adjust the constant term in (29) by setting $x = 0$. The expansion is valid for all values of x.

6. Recourse to a differential equation

To obtain the expansions of powers of $\sin^{-1}x$ the foregoing methods must be somewhat extended. Let us consider first the expansion of $(\sin^{-1}x)^2$. We write

$$\frac{\sin^{-1}x}{\sqrt{(1-x^2)}} = a_1 x+a_3 x^3+a_5 x^5+\ldots \quad \text{to } \infty. \tag{30}$$

We may assume the series of odd powers only, on the right, since we know that the expression on the left is an odd function of x. Then

$$\sqrt{(1-x^2)}\sin^{-1}x = a_1 x+(a_3-a_1)x^3+(a_5-a_3)x^5+\ldots.$$

Differentiation gives

$$1-\frac{x\sin^{-1}x}{\sqrt{(1-x^2)}} = a_1+3(a_3-a_1)x^2+5(a_5-a_3)x^4+\ldots,$$

and so, from (30),

$$1 = a_1+(3a_3-2a_1)x^2+(5a_5-4a_3)x^4+\ldots.$$

Hence we must have

$$a_1 = 1, \qquad 3a_3 = 2a_1, \qquad 5a_5 = 4a_3, \qquad \ldots,$$

and so (30) is

$$\frac{\sin^{-1}x}{\sqrt{(1-x^2)}} = x+\frac{2}{3}x^3+\frac{2.4}{3.5}x^5+\ldots \quad \text{to } \infty. \tag{31}$$

Integration gives

$$\tfrac{1}{2}(\sin^{-1}x)^2 = \frac{x^2}{2}+\frac{2}{3}.\frac{x^4}{4}+\frac{2.4}{3.5}.\frac{x^6}{6}+\ldots \quad \text{to } \infty. \tag{32}$$

Of these two expansions, (31) converges in the open interval $|x| < 1$ and (32) in the closed interval $|x| \leqslant 1$.

We may extend formulae (31), (32) to higher powers of $\sin^{-1}x$. Thus, write

$$\frac{(\sin^{-1}x)^2}{2\sqrt{(1-x^2)}} \equiv \frac{1}{2}a_2x^2+\frac{1.3}{2.4}a_4x^4+\frac{1.3.5}{2.4.6}a_6x^6+\dots \quad \text{to } \infty,$$

the expression on the left being clearly an even function. Then, as before,

$$\tfrac{1}{2}(\sin^{-1}x)^2\sqrt{(1-x^2)} = a_2\frac{x^2}{2}+\frac{1}{2}(3a_4-4a_2)\frac{x^4}{4}+\frac{1.3}{2.4}(5a_6-6a_4)\frac{x^6}{6}+\dots.$$

Differentiation gives

$$\sin^{-1}x-\frac{x(\sin^{-1}x)^2}{2\sqrt{(1-x^2)}} = a_2x+\frac{1}{2}(3a_4-4a_2)x^3+\frac{1.3}{2.4}(5a_6-6a_4)x^5+\dots,$$

i.e.

$$x+\frac{1}{2}.\frac{x^3}{3}+\frac{1.3}{2.4}.\frac{x^5}{5}+\dots = a_2x+\frac{1}{2}(3a_4-3a_2)x^3+\frac{1.3}{2.4}(5a_6-5a_4)x^5+\dots.$$

Thus $\qquad a_2 = 1, \qquad a_4-a_2 = \frac{1}{3^2}, \qquad a_6-a_4 = \frac{1}{5^2}, \qquad \dots,$

which gives

$$\frac{(\sin^{-1}x)^2}{2!\sqrt{(1-x^2)}} = \frac{1}{2}x^2+\left(1+\frac{1}{3^2}\right)\frac{1.3}{2.4}x^4+\left(1+\frac{1}{3^2}+\frac{1}{5^2}\right)\frac{1.3.5}{2.4.6}x^6+\dots \quad \text{to } \infty, \tag{33}$$

$$\frac{(\sin^{-1}x)^3}{3!} = \frac{1}{2}.\frac{x^3}{3}+\left(1+\frac{1}{3^2}\right)\frac{1.3}{2.4}.\frac{x^5}{5}+\left(1+\frac{1}{3^2}+\frac{1}{5^2}\right)\frac{1.3.5}{2.4.6}.\frac{x^6}{6}+\dots \quad \text{to } \infty. \tag{34}$$

The general formulae can be shown to be

$$\begin{aligned}\frac{(\sin^{-1}x)^{2n}}{(2n)!} = {} & C_{2n-2,n-1}\frac{2.4\dots(2n-2)}{3.5\dots(2n-1)}.\frac{x^{2n}}{2n}+\\ & +C_{2n,n-1}\frac{2.4\dots2n}{3.5\dots(2n+1)}.\frac{x^{2n+2}}{2n+2}+\\ & +C_{2n+2,n-1}\frac{2.4\dots(2n+2)}{3.5\dots(2n+3)}.\frac{x^{2n+4}}{2n+4}+\dots \quad \text{to } \infty,\end{aligned} \tag{35}$$

$$\begin{aligned}\frac{(\sin^{-1}x)^{2n+1}}{(2n+1)!} = {} & C_{2n-1,n}\frac{1.3\dots(2n-1)}{2.4\dots2n}.\frac{x^{2n+1}}{2n+1}+\\ & +C_{2n+1,n}\frac{1.3\dots(2n+1)}{2.4\dots(2n+2)}.\frac{x^{2n+3}}{2n+3}+\\ & +C_{2n+3,n}\frac{1.3\dots(2n+3)}{2.4\dots(2n+4)}.\frac{x^{2n+5}}{2n+5}+\dots \quad \text{to } \infty,\end{aligned} \tag{36}$$

where $C_{2n,r}$ and $C_{2n+1,r}$ denote the sums of the products, r at a time, of 2^{-2}, 4^{-2},..., $(2n)^{-2}$ and 1, 3^{-2},..., $(2n+1)^{-2}$ respectively. We give

subsequently another proof of these results that is more easily remembered.

The corresponding formulae for powers of $\sinh^{-1}x$, i.e. of

$$\log\{x+\sqrt{(1+x^2)}\},$$

can be written down by analogy.

The argument underlying the above analysis has been tacitly this: we know *a priori* that $(1-x^2)^{-\frac{1}{2}}\sin^{-1}x$ is analytic and can be represented by a power-series of the form (30); this power-series and its first derivative satisfy a certain identity and, as a consequence, its coefficients must be of the form stated. We can, however, recast the argument so as to dispense with the presumption of expansibility. The identity satisfied, for instance, by the function

$$y = (1-x^2)^{-\frac{1}{2}}\sin^{-1}x$$

is actually

$$D\{y(1-x^2)\} = 1-xy,$$

i.e. is the differential equation

$$(1-x^2)\frac{dy}{dx}-xy = 1. \tag{37}$$

By retracing our analysis we can show that the series (31) also satisfies this differential equation. Thus the series and the given function are both solutions of the same differential equation. We have still to show that they are the same solution. Now the differential equation is linear and of the first order. Its general solution is therefore of the form

$$y = u+Av,$$

where u, v are given functions and A is the arbitrary constant. Actually, the general solution is

$$y = (1-x^2)^{-\frac{1}{2}}\sin^{-1}x+A(1-x^2)^{-\frac{1}{2}}.$$

Thus, if two solutions of the equation are equal for one value of x that is not also a zero of v, e.g. $x = 0$, then they are equal for all values of x within their common domain of definition.

This is the method that we now develop: the formation of a differential equation satisfied by the given function; the solution of the differential equation by a power-series; the identification of these two solutions of the differential equation.

We may, if we please, regard the method of the last two sections as only a particular case of this method, if we think of it as based on the formation and solution of a differential equation of the rudimentary form

$$\frac{dy}{dx} = f(x).$$

7. Approximate expansions

This method can also be applied to obtain an approximate or partial expansion when the complete expansion is unobtainable, that is to say, an expansion of the type considered in chapter VII § 9. For example, if x is sufficiently small, we may assume, since $\tan x$ is an odd function, the approximate expansion

$$\tan x = a_1 x + a_3 x^3 + a_5 x^5 + \dots .$$

By differentiation,

$$\sec^2 x = a_1 + 3a_3 x^2 + 5a_5 x^4 + \dots,$$

which gives the identity

$$a_1 + 3a_3 x^2 + 5a_5 x^4 + \dots = 1 + (a_1 x + a_3 x^3 + a_5 x^5 + \dots)^2.$$

Equating coefficients we obtain

$$a_1 = 1, \qquad 3a_3 = a_1^2, \qquad 5a_5 = 2a_1 a_3,$$
$$7a_7 = 2a_1 a_5 + a_3^2, \qquad 9a_9 = 2a_1 a_7 + 2a_3 a_5,$$
$$11a_{11} = 2a_1 a_9 + 2a_3 a_7 + a_5^2, \qquad \dots .$$

Solving these equations in succession we get the expansion

$$\tan x = x + \frac{x^3}{3} + \frac{2x^5}{3.5} + \frac{17x^7}{5.7.9} + \frac{62x^9}{5.7.9^2} + \frac{1382x^{11}}{5^2.7.9^2.11} + \dots . \tag{38}$$

Similarly for the odd function $\cot x$ we assume the expansion

$$\cot x = x^{-1} - a_1 x - a_3 x^3 - a_5 x^5 - \dots .$$

Differentiation gives the identity

$$x^{-2} + a_1 + 3a_3 x^2 + 5a_5 x^4 + \dots = 1 + (x^{-1} - a_1 x - a_3 x^3 - a_5 x^5 - \dots)^2,$$

whence we have

$$a_1 = 1 - 2a_1, \qquad 3a_3 = -2a_3 + a_1^2, \qquad 5a_5 = -2a_5 + 2a_1 a_3,$$
$$7a_7 = -2a_7 + 2a_1 a_5 + a_3^2, \qquad 9a_9 = -2a_9 + 2a_1 a_7 + 2a_3 a_5, \qquad \dots .$$

Solving, we obtain the expansion

$$\cot x = x^{-1} - \frac{x}{3} - \frac{x^3}{3^2.5} - \frac{2x^5}{3.5.7.9} - \frac{x^7}{3.5^2.7.9} - \frac{2x^9}{3.5.7.9^2.11} - \dots . \tag{39}$$

From these partial expansions of $\tan x$, $\cot x$ we can similarly obtain the partial expansions of $\sec x$, $\operatorname{cosec} x$. For assume

$$\sec x = 1 + a_2 x^2 + a_4 x^4 + a_6 x^6 + \dots .$$

Then, by differentiation,

$$\sec x \tan x = 2a_2 x + 4a_4 x^3 + 6a_6 x^5 + \dots$$

and, so from (38),

$$2a_2x+4a_4x^3+6a_6x^5+\dots$$
$$=(1+a_2x^2+a_4x^4+a_6x^6+\dots)(x+\tfrac{1}{3}x^3+\tfrac{2}{15}x^5+\tfrac{17}{315}x^7+\dots).$$

Thus
$$2a_2=1,\qquad 4a_4=\tfrac{1}{3}+a_2,\qquad 6a_6=\tfrac{2}{15}+\tfrac{1}{3}a_2+a_4,$$
$$8a_8=\tfrac{17}{315}+\tfrac{2}{15}a_2+\tfrac{1}{3}a_4+a_6,\qquad \dots,$$

which give
$$\sec x=1+\frac{1}{2!}x^2+\frac{5}{4!}x^4+\frac{61}{6!}x^6+\frac{1385}{8!}x^8+\dots. \tag{40}$$

This method is no neater than that of dividing unity by the cosine series, but serves as a useful check on the result so obtained.

The corresponding series for $\operatorname{cosec} x$ is

$$\operatorname{cosec} x=x^{-1}+\frac{x}{3!}+\frac{7x^3}{3.5!}+\frac{31x^5}{3.7!}+\dots, \tag{41}$$

which we may obtain either by differentiation, by division of the sine series into unity, or from either of the identities

$$\operatorname{cosec} x=\tan\tfrac{1}{2}x+\cot x=\cot\tfrac{1}{2}x-\cot x.$$

8. Hypergeometric functions

We consider now the general application of this method in which the differential equations employed are no longer necessarily of the first order. Evidently, the class of functions to which the method applies are those that appear as solutions of the class of differential equations that are readily solved in power-series. We can give the most systematic account of the theory, if we approach it from the point of view of the differential equations themselves and the determination of their solutions in power-series. The reason is that only in certain cases and, as it were, accidentally, is the solution of a differential equation of this class an elementary function. For the most part these differential equations define new transcendental functions by their solution: in fact, many of the important functions of analysis first appear as solutions of such equations. In starting, then, from the differential equations we take the more comprehensive, and therefore the more systematic, view.

Now the class of differential equations to which we have been referring can be written, in the notation of the preceding chapter, in the form

$$f(\delta)y=x^m g(\delta)y, \tag{42}$$

where f, g denote polynomial functions of their arguments and m is any

number. If we take x^m as a new independent variable, the differential equation reduces to the simpler form

$$f(\delta)y = x\,g(\delta)y, \tag{43}$$

in which form we now consider it.

The differential equation (43) is usually known as a *generalized hypergeometric* differential equation, the corresponding equation of the second order being the hypergeometric equation in the strict sense. If, however, we are not always punctilious in inserting 'generalized', little harm is done and the title gains in brevity.

Now define the terminated power-series

$$S_n(a) = x^a + \frac{g(a)}{f(a+1)}x^{a+1} + \frac{g(a)g(a+1)}{f(a+1)f(a+2)}x^{a+2} + \dots + \frac{g(a)g(a+1)\dots g(a+n-1)}{f(a+1)f(a+2)\dots f(a+n)}x^{a+n}. \tag{44}$$

Then, in virtue of the fundamental property† of the operator δ, we have

$$f(\delta)\,S_n(a) = f(a)x^a + g(a)x^{a+1} + \frac{g(a)g(a+1)}{f(a+1)}x^{a+2} + \dots + \frac{g(a)g(a+1)\dots g(a+n-1)}{f(a+1)\dots f(a+n-1)}x^{a+n},$$

$$x\,g(\delta)\,S_n(a) = g(a)\,x^{a+1} + \frac{g(a)g(a+1)}{f(a+1)}x^{a+2} + \frac{g(a)g(a+1)g(a+2)}{f(a+1)f(a+2)}x^{a+3} + \dots + \frac{g(a)g(a+1)\dots g(a+n)}{f(a+1)\dots f(a+n)}x^{a+n+1}.$$

Thus

$$\{f(\delta) - x\,g(\delta)\}S_n(a) = f(a)\,x^a - \frac{g(a)g(a+1)\dots g(a+n)}{f(a+1)\dots f(a+n)}x^{a+n+1}, \tag{45}$$

the intervening terms cancelling obliquely.

More generally, if $S(a)$ denote the corresponding infinite series

$$S(a) \equiv x^a + \frac{g(a)}{f(a+1)}x^{a+1} + \frac{g(a)g(a+1)}{f(a+1)f(a+2)}x^{a+2} + \dots \quad \text{to } \infty, \tag{46}$$

supposed convergent, then we have similarly

$$\{f(\delta) - x\,g(\delta)\}S(a) = f(a)\,x^a, \tag{47}$$

the differentiations being permissible within the interval of convergence. In particular, if $f(a) = 0$, then

$$\{f(\delta) - x\,g(\delta)\}S(a) = 0,$$

† Chapter XIII (30).

that is to say,

(48) *If a is a root of the equation $f(a) = 0$, then*

$$S(a) = x^a + \frac{g(a)}{f(a+1)} x^{a+1} + \frac{g(a)\,g(a+1)}{f(a+1)\,f(a+2)} x^{a+2} + \dots \quad \text{to } \infty,$$

supposed convergent, is a solution of the hypergeometric equation

$$f(\delta)y = x\,g(\delta)y.$$

The equation $f(a) = 0$ is known as the *indicial equation*, since it determines the index of the leading term of $S(a)$. Each root, then, of the indicial equation determines a corresponding formal solution $S(a)$ and we have, in general, a number of solutions equal to the degree of the indicial equation.

These series $S(a)$ are series expressed in *ascending* powers of x. But we can determine similar solutions in descending powers of x. For, writing (43) in the form

$$g(\delta)y = x^{-1} f(\delta)y, \tag{49}$$

and defining the terminated series

$$T_n(b) = x^b + \frac{f(b)}{g(b-1)} x^{b-1} + \frac{f(b)f(b-1)}{g(b-1)\,g(b-2)} x^{b-2} + \dots +$$
$$+ \frac{f(b)f(b-1)\dots f(b-n+1)}{g(b-1)\,g(b-2)\dots g(b-n)} x^{b-n} \tag{50}$$

and the corresponding infinite series

$$T(b) = x^b + \frac{f(b)}{g(b-1)} x^{b-1} + \frac{f(b)f(b-1)}{g(b-1)\,g(b-2)} x^{b-2} + \dots \quad \text{to } \infty, \tag{51}$$

we obtain, in analogy with (45), (47),

$$\{g(\delta) - x^{-1} f(\delta)\} T_n(b) = g(b)x^b - \frac{f(b)f(b-1)\dots f(b-n)}{g(b-1)\dots g(b-n)} x^{b-n-1} \tag{52}$$

and

$$\{g(\delta) - x^{-1} f(\delta)\} T(b) = g(b)x^b. \tag{53}$$

Thus again, $y = T(b)$ is a solution of (49), if b satisfies the algebraic equation

$$g(b) = 0, \tag{54}$$

which we therefore take as the indicial equation of the descending series $T(b)$. Here, too, we obtain formal solutions of the differential equation equal in number to the degree of the indicial equation (54).

9. Convergence of the series

We have now to see how far the formal solutions $S(a)$, $T(b)$ provide actual solutions of the differential equations, and, in particular, how far

the complete solution of the differential equation can be expressed in terms of them. These formal solutions may fail in any of three ways:

(α) a factor $f(a+r)$ or $g(b-s)$ in the denominator of $S(a)$ or $T(b)$ may be zero, and the series therefore be meaningless from that term onwards;

(β) for certain values or for a certain range of values of x, the series may not converge;

(γ) either indicial equation may have repeated roots, in which case the corresponding series $S(a)$ or $T(b)$ will not be all distinct.

We discuss these possibilities in the order (β), (γ), (α).

Now, if u_N denote the Nth term in the series $S(a)$, we have

$$\frac{u_{N+1}}{u_N} = \frac{g(a+N-1)}{f(a+N)}x.$$

Let the leading terms of the polynomials $f(\delta)$, $g(\delta)$ be respectively

$$f(\delta) \equiv p\delta^m+p_1\delta^{m-1}+\dots, \qquad g(\delta) \equiv q\delta^n+q_1\delta^{n-1}+\dots.$$

Then $$\frac{u_{N+1}}{u_N} = \frac{qN^n+\{qn(a-1)+q_1\}N^{n-1}+O(N^{n-2})}{pN^m+\{pma+p_1\}N^{m-1}+O(N^{m-2})}x, \tag{55}$$

and so as $N \to \infty$, $$\left|\frac{u_{N+1}}{u_N}\right| \to \begin{cases} 0 & (m > n), \\ |qx/p| & (m = n), \\ \infty & (m < n). \end{cases}$$

Hence† the series $S(a)$ converges

(i) for every (finite) x, if $m > n$;

(ii) in the open‡ interval $|x| < |p/q|$, if $m = n$;

(iii) at $x = 0$ alone, if $m < n$.

Thus in (iii) the formal solution $S(a)$ is valueless except possibly as an asymptotic expansion.

By interchange of f and g, and, at the same time, of x and x^{-1}, we find similarly that the series $T(b)$ converges

(i) for every x (except $x = 0$), if $m < n$;

(ii) in the region $|x| > |p/q|$, if $m = n$;

(iii) at '$x = \infty$' alone, if $m > n$.

Again we must exclude (iii) except perhaps for asymptotic purposes.

† By the fundamental 'ratio test' presumably devised by Gauss for the hypergeometric series itself: cf. Bromwich, *Infinite Series* (1926), 39-41 (§ 12.2).

‡ Actually, if $m = n$, the series has the interval of convergence $(-p/q, p/q)$. We discuss later whether it will be open or closed: at present we exclude the end-points from the discussion.

Since m, n are the respective degrees of the indicial equations $f(a) = 0$, $g(b) = 0$, we have now proved, subject to the overriding conditions (α), (γ) above, that

(i) if $m > n$, we have m solutions $S(a)$, convergent for every (finite) x;

(ii) if $m < n$, we have n solutions $T(b)$, convergent for every x (except $x = 0$);

(iii) if $m = n$, we have m solutions $S(a)$ which converge inside the open interval $)-p/q, p/q($ and m solutions $T(b)$ which converge outside the same interval closed, so that there is convergence of one or other type of series at all points except possibly the two points $x = \pm p/q$.

Now the order of the differential equation (43) is the order of $f(\delta)$ or $g(\delta)$, whichever is the greater, i.e. is the greater of m, n (or either of them, if they are equal). Moreover, by theorem (20) of chapter V, the complete solution of a linear differential equation of order n is a linear function of any n particular solutions that are themselves linearly distinct. Thus, we shall have determined the complete solution of the differential equation (43), when we have obtained m or n distinct solutions, whichever is the greater number.

It follows, then, that, still subject to the overriding conditions (α), (γ), we have obtained the complete solution of the differential equation in ascending series $S(a)$, if $m > n$; and in descending series $T(b)$, if $m < n$. In the latter case we have no information at the special point $x = 0$. If $m = n$, we have the complete solution of the equation except at the two points $x = \pm p/q$: between these points in ascending series $S(a)$; outside these points in descending series $T(b)$.

In the case $m = n$ the leading term in the differential equation is

$$(p-qx)\delta^m y,$$

and the excluded point $x = p/q$ is accordingly a zero of the coefficient of this term. By the general theory of linear differential equations such a point is ordinarily a singularity of the solution of the differential equation:† the excluded points $x = \infty$, $x = 0$ in the cases $m \gtrless n$ fall into the same category. The other excluded point $x = -p/q$ occurs as the image in the origin of the singular point $x = p/q$.

Since here we are going to use these methods only for the expansion of known functions, we need not further discuss the theory of these singularities, which can be recognized from the known behaviour of the functions themselves.

† In theorem (20) of chapter V, already quoted, zeros of the leading coefficient of the differential equation are expressly excluded from the region in which the general solution is valid.

10. Convergence at the end-points

In the case $m = n$ the points $x = \pm p/q$ have so far only been excluded from our discussion: we have not actually proved non-convergence there. Let us now discuss the convergence of the ascending series $S(a)$ at these points, which are the end-points of its interval of convergence.

At $x = p/q$ we have from (55)

$$\frac{u_{N+1}}{u_N} = \frac{N^n+\{n(a-1)+q_1/q\}N^{n-1}+O(N^{n-2})}{N^n+\{na+p_1/p\}N^{n-1}+O(N^{n-2})}$$
$$= 1-(p_1/p+n-q_1/q)N^{-1}+O(N^{-2}).$$

Now the 'critical' ratio is

$$u_{N+1}/u_N = 1-N^{-1}-(N\log N)^{-1}-\dots.$$

Thus $S(a)$ converges or diverges at $x = p/q$, according as

$$n+p_1/p-q_1/q > 1 \quad \text{or} \quad < 1. \tag{56}$$

At $x = -p/q$ we have similarly

$$-u_{N+1}/u_N = 1-(p_1/p+n-q_1/q)N^{-1}+O(N^{-2}). \tag{57}$$

Thus $$\lim(u_{N+1}/u_N) = -1,$$

and therefore, ultimately, the terms alternate in sign. Again from (57)

$$\log|u_{N+1}|-\log|u_N| = -(p_1/p+n-q_1/q)N^{-1}+O(N^{-2})$$

Hence, on summation,

$$\log|u_{N+1}| = -(p_1/p+n-q_1/q)\log N+O(1),$$

and so $$|u_{N+1}| \sim N^{-(p_1/p+n-q_1/q)}.$$

Accordingly the condition

$$p_1/p+n-q_1/q > 0 \tag{58}$$

is necessary to secure $u_N \to 0$, and therefore to secure convergence of the series. Moreover, if (58) is satisfied, then, by (57), we have ultimately $|u_{N+1}/u_N| < 1$. Hence, if (58) is satisfied, the series is ultimately an alternating series of terms that converge monotonically to zero, and is therefore a convergent series. In other words, (58) is both necessary and sufficient for convergence.

We obtain corresponding conditions for the convergence of the descending series $T(b)$ at the same end-points $x = \pm p/q$.

The facts which we have thus far obtained about the solution of the differential equation may be summed up in the following enunciation:

(59) *The complete solution of the differential equation*

$$f(\delta)y = xg(\delta)y$$

can, in general, be exhibited in terms of the series

$$S(a) = x^a + \frac{g(a)}{f(a+1)}x^{a+1} + \frac{g(a)g(a+1)}{f(a+1)f(a+2)}x^{a+2} + \dots \quad \textit{to } \infty,$$

$$T(b) = x^b + \frac{f(b)}{g(b-1)}x^{b-1} + \frac{f(b)f(b-1)}{g(b-1)g(b-2)}x^{b-2} + \dots \quad \textit{to } \infty.$$

If the order of $f(\delta)$ exceeds that of $g(\delta)$, then the complete solution is, in general, given by the ascending series $S(a)$, where a runs through the zeros of $f(a)$; and these series converge for every (finite) x.

If $g(\delta)$ is of higher order than $f(\delta)$, then the complete solution is, in general, given by the descending series $T(b)$, where b runs through the zeros of $g(b)$; and these series converge everywhere (except at $x = 0$).

If $f(\delta)$, $g(\delta)$ are of equal order, then the complete solution is, in general, given inside the interval $(-p/q, p/q)$ by the ascending (convergent) series $S(a)$, where a runs through the zeros of $f(a)$, and outside this interval by the descending (convergent) series $T(b)$, where b runs through the zeros of $g(b)$.

If $\qquad f(\delta) = p\delta^n + p_1\delta^{n-1} + \dots, \qquad g(\delta) = q\delta^n + q_1\delta^{n-1} + \dots,$

then the necessary and sufficient conditions that

$S(a)$, $\qquad T(b) \qquad$ *converge at* p/q, $\qquad -p/q$,

are that

$p_1/p - q_1/q + n, \qquad q_1/q - p_1/p + n \qquad > 1, \qquad > 0.$

The exceptional cases in the above theorem covered by the qualifying 'in general' are the following:

(i) if the indicial equation $f = 0$ has negative roots, there will be terms of negative index in the corresponding ascending series, and some obvious modification will have to be made to retain convergence at $x = 0$;

(ii) if the roots of an indicial equation are not all real, the corresponding series will contain complex indices and coefficients: although these series, taken in conjugate pairs, can still be expressed in a real form, we shall exclude such cases as beyond our present scope;

(iii) there still remain the overriding conditions (α), (γ) above to which we now turn our attention.

11. Repeated roots of the indicial equation

Of the possibilities (α), (γ) now to be discussed we take first (γ), namely the possibility that the indicial equation $f(a) = 0$ have repeated roots.

Returning to the equation (47), on which the whole of the foregoing discussion has been based, we regard a still as an arbitrary parameter and differentiate partially in a. This gives

$$\{f(\delta)-x\,g(\delta)\}S'(a) = f(a)\,x^a\log x+f'(a)\,x^a. \tag{60}$$

On the left we have differentiated in a across the differentiations δ, that is to say, we have supposed the operations $\partial/\partial x$, $\partial/\partial a$ to be commutative. This we know† to be the case, if the resulting expression is continuous. On the left again we shall mean by $S'(a)$ the term-by-term derivative of $S(a)$ with respect to a. This term-by-term differentiability of $S(a)$ and its continuity in a are both guaranteed, if the series $S(a)$ converges uniformly in a, x jointly. It is better, however, to postpone the question of the validity of (60) and consider, for the present, only its formal aspect.

If a is now a repeated root of $f(a) = 0$, then $f(a) = 0 = f'(a)$, and so, by (60), $S'(a)$ is a solution of

$$\{f(\delta)-x\,g(\delta)\}y = 0.$$

Thus $S'(a)$ replaces the solution which has been lost by the coalescence of two roots of the indicial equation, and the two solutions, corresponding to the double root a of the indicial equation, are $S(a)$, $S'(a)$. So generally, if the indicial equation has n repeated roots a, the corresponding n solutions of the differential equation are

$$S(a),\, S'(a),\ldots,\, S_{n-1}(a),$$

as we see by further partial differentiation of (60).

We may notice incidentally that, if a is a simple root of the indicial equation, then $S'(a)$ is a solution of the extended equation

$$\{f(\delta)-x\,g(\delta)\}y = f'(a)x^a,$$

and so on.

To obtain an explicit expression for $S'(a)$ write now

$$S(a) \equiv x^a+A_1x^{a+1}+A_2x^{a+2}+\ldots \quad \text{to } \infty,$$

where a is once more to be regarded as an arbitrary parameter. Then, so long as we may differentiate term by term, we have

$$S'(a) = \log x\{x^a+A_1x^{a+1}+A_2x^{a+2}+\ldots\}+\frac{\partial A_1}{\partial a}x^{a+1}+\frac{\partial A_2}{\partial a}x^{a+2}+\ldots.$$

Now
$$A_N = \frac{g(a)\,g(a+1)\ldots g(a+N-1)}{f(a+1)f(a+2)\ldots f(a+N)},$$

and so
$$\frac{\partial A_N}{\partial a} = A_N\sum_{r=1}^{N}\left\{\frac{g'(a+r-1)}{g(a+r-1)}-\frac{f'(a+r)}{f(a+r)}\right\}.$$

† Chapter VI (27).

Again, suppose that

$$f(\delta) \equiv p\prod_{s=1}^{m}(\delta-a_s), \qquad g(\delta) \equiv q\prod_{t=1}^{n}(\delta-b_t).$$

Then
$$\frac{\partial A_N}{\partial a} = -A_N\sum_{r=1}^{N}\left(\sum_{s=1}^{m}\frac{1}{a-a_s+r}-\sum_{t=1}^{n}\frac{1}{a-b_t+r-1}\right). \tag{61}$$

But f, by hypothesis, has the double zero a: we may therefore take $a_1 = a = a_2$, and so at length obtain

$$S'(a) = \log x\; S(a)-\sum_{N=1}^{\infty}x^{a+N}A_N\sum_{r=1}^{N}\left(\frac{2}{r}+\sum_{s=3}^{m}\frac{1}{a-a_s+r}-\sum_{t=1}^{n}\frac{1}{a-b_t+r-1}\right). \tag{62}$$

The extra solution corresponding to the additional root a of the indicial equation accordingly consists of two parts: the original solution $S(a)$ multiplied by $\log x$, and a new series in which the coefficients of the original series are multiplied by sums of 'harmonic' series

$$\sum_{r=1}^{N}(c+r)^{-1}.$$

We should obtain in similar fashion

$$S''(a) = (\log x)^2S(a)+2\log x\sum_{N=1}^{\infty}\frac{\partial A_N}{\partial a}x^{a+N}+\sum_{N=1}^{\infty}\frac{\partial^2 A_N}{\partial a^2}x^{a+N},$$

and so on for further partial differentiations in a. To complete the explicit expression of $S^{(k)}(a)$ we need a convenient expression for $\partial^k A_N/\partial a^k$. Let us write, for any positive integer k,

$$C_{Nk} = \sum_{r=1}^{N}\left\{\sum_{s=1}^{m}(a-a_s+r)^{-k}-\sum_{t=1}^{n}(a-b_t+r-1)^{-k}\right\}. \tag{63}$$

Then
$$\frac{\partial C_{Nk}}{\partial a} = -kC_{N,k+1}.$$

Now we have just proved that

$$\frac{\partial A_N}{\partial a} = -C_{N1}A_N,$$

and accordingly

$$\frac{\partial^2 A_N}{\partial a^2} = -C_{N1}\frac{\partial A_N}{\partial a}+C_{N2}A_N = (C_{N1}^2+C_{N2})A_N.$$

We similarly obtain

$$\frac{\partial^3 A_N}{\partial a^3} = (C_{N1}^3+3C_{N1}C_{N2}+2C_{N3})A_N,$$

$$\frac{\partial^4 A_N}{\partial a^4} = (C_{N1}^4+6C_{N1}^2C_{N2}+8C_{N1}C_{N3}+3C_{N2}^2+6C_{N4})A_N$$

and so forth. We can use these results to give the following compact expressions for $S'(a)$, $S''(a)$, etc.:

$$\left.\begin{aligned}
S'(a) &= \sum_{N=0}^{\infty} (\log x - C_{N1}) A_N x^{a+N}, \\
S''(a) &= \sum_{N=0}^{\infty} \{(\log x - C_{N1})^2 + C_{N2}\} A_N x^{a+N}, \\
S'''(a) &= \sum_{N=0}^{\infty} \{(\log x - C_{N1})^3 + 3(\log x - C_{N1}) C_{N2} - 2C_{N3}\} A_N x^{a+N}, \\
S^{(4)}(a) &= \sum_{N=0}^{\infty} \{(\log x - C_{N1})^4 + 6C_{N2}(\log x - C_{N1})^2 - \\
&\qquad - 8C_{N3}(\log x - C_{N1}) + 3(C_{N2}^2 + 2C_{N4})\} A_N x^{a+N},
\end{aligned}\right\} \tag{64}$$

with the conventions $A_0 = 1$, $C_{01} = 0$.

We can deal in similar fashion with repeated roots of the indicial equation $g(b) = 0$.

12. Roots of the indicial equation differing by an integer

We come at last to the possibility (α), namely that the series $S(a)$ fail because of a zero factor $f(a+n)$ in its denominator. In such a case, then, the indicial equation $f = 0$ has roots a, $a+n$ differing by an integer n. This, in a sense, may be regarded as a generalization of the case discussed in the preceding section in which two roots of the indicial equation differed by zero: more pertinently we proceed to show how its discussion can be based on the results obtained for this case of equal roots.

We may observe at the start that these two roots a, $a+n$ of the indicial equation give rise to two formal solutions $S(a)$, $S(a+n)$, and that of these two solutions it is only $S(a)$ that is open to objection on the score of vanishing factors in the denominators of the coefficients: the solution $S(a+n)$ is, in general, unexceptionable and so stands.

Let us pick out the two relevant factors of f and write

$$f(\delta) = (\delta - a)(\delta - a - n)\, F(\delta).$$

In the differential equation

$$(\delta - a)(\delta - a - n)\, F(\delta) y = x g(\delta) y \tag{65}$$

set

$$y = (\delta - a)(\delta - a - 1)\dots(\delta - a - n + 1) z, \tag{66}$$

getting, after rearrangement of δ-factors,

$$\begin{aligned}
(\delta - a - 1)\dots(\delta - a - n)(\delta - a)^2\, F(\delta) z &\\
= x(\delta - a)(\delta - a - 1)\dots(\delta - a - n + 1)\, g(\delta) z &\\
= (\delta - a - 1)\dots(\delta - a - n)\, x g(\delta) z. &
\end{aligned}$$

This equation is certainly satisfied, if it is satisfied after removal of the operator $(\delta-a-1)...(\delta-a-n)$, i.e. if z satisfies the equation

$$(\delta-a)^2 F(\delta)z = xg(\delta)z. \tag{67}$$

This is a differential equation whose indicial equation has equal roots, and the theory of the preceding section applies. Write $\bar{S}(a)$, $\bar{S}'(a)$ for the pair of solutions of (67) associated with the repeated root a of the indicial equation. We then have, for (65), the corresponding solutions

$$(\delta-a)(\delta-a-1)...(\delta-a-n+1)\,\bar{S}(a), \tag{68}$$

$$(\delta-a)(\delta-a-1)...(\delta-a-n+1)\,\bar{S}'(a). \tag{69}$$

Now $\bar{S}(a)$ begins with the term x^a. But the operator

$$(\delta-a)...(\delta-a-n+1)$$

removes every term before x^{a+n}. Thus the series (68) begins with x^{a+n} and is evidently a constant multiple of the series $S(a+n)$. We need, then, preserve only (69) to take the place of the missing $S(a)$. It is best preserved in its symbolic form, being worked out explicitly for each particular case.

Similarly, if the indicial equation $f = 0$ has *three* roots $a, a+m, a+n$, where m, n are positive integers, we make the substitution

$$y = \prod_{r=0}^{m-1}(\delta-a-r)\prod_{s=0}^{n-1}(\delta-a-s)z = \Delta z, \text{ say},$$

and obtain three solutions in the form

$$\Delta\,\bar{S}(a), \qquad \Delta\,\bar{S}'(a), \qquad \Delta\,\bar{S}''(a).$$

It should be remarked that in this case m, n need not be unequal.

We can proceed similarly when more than three roots of $f = 0$ differ by integers, and so too when roots of $g = 0$ differ by integers.

We may appropriately consider here the possibility that a root of $g = 0$ *exceed* by a *positive* integer or zero a root of $f = 0$, i.e. that simultaneously

$$f(a) = 0, \qquad g(a+n) = 0 \qquad (n \geqslant 0).$$

In this case the coefficients of $S(a)$ vanish after the term in x^{a+n} and the series in this case *terminates*. The same terminated series† $S(a)$ appears, of course, as a terminated descending series $T(a+n)$, being now terminated by the zero factor $f(a)$ in the numerator.

It may even happen that a root of $g = 0$ is interposed between two consecutive roots of $f = 0$, all the differences being integers, i.e. that simultaneously

$$f(a) = 0, \qquad g(a+m) = 0, \qquad f(a+m+n) = 0 \quad (m \geqslant 0;\ n > 0).$$

† Or, more precisely, some constant multiple of it.

In such a case the series $S(a)$ terminates at x^{a+m} before the zero factor $f(a+m+n)$ appears in the denominators of the coefficients. In other words, $S(a)$ is now unexceptionable as a solution, and the two solutions associated with the roots a, $a+m+n$ of $f=0$ are once again simply

$$S(a), \qquad S(a+m+n).$$

13. Justification of the formal operations

The analysis of the two preceding sections has, so far, merely a formal significance; for it is based on equation (60)

$$\{f(\delta)-x\,g(\delta)\}S'(a)=f(a)x^a\log x+f'(a)x^a,$$

in which we have differentiated in a across the operator $f(\delta)-x\,g(\delta)$, and we have, moreover, taken $S'(a)$ as meaning the series arising from term-by-term differentiation of $S(a)$. As we have said, these formal operations and the further formal operations involving $S''(a)$, $S'''(a)$,... can be justified by establishing over some suitable region of a, x the uniform convergence of $S'(a)$, $S''(a)$, $S'''(a)$,... .

The actual discussion of the uniform convergence I find tiresome, and I prefer to give an alternative method of arriving at (60), in which only the ordinary convergence of the series need be considered.

Now
$$S'(a)=\log x\;S(a)+\sum_{N=1}^{\infty}\frac{\partial A_N}{\partial a}x^{a+N}$$

and, by chapter XIII (37),

$$\begin{aligned}\{f(\delta)-x\,g(\delta)\}&\log x\;S(a)\\&=\log x\{f(\delta)-x\,g(\delta)\}S(a)+\{f'(\delta)-x\,g'(\delta)\}S(a)\\&=f(a)x^a\log x+\{f'(\delta)-x\,g'(\delta)\}S(a),\quad\text{by (47).}\end{aligned}$$

Thus

$$\begin{aligned}&\{f(\delta)-x\,g(\delta)\}\;S'(a)\\&=f(a)x^a\log x+\{f'(\delta)-x\,g'(\delta)\}\sum_{N=0}^{\infty}A_N x^{a+N}+\{f(\delta)-x\,g(\delta)\}\sum_{N=1}^{\infty}\frac{\partial A_N}{\partial a}x^{a+N}\\&=f(a)x^a\log x+f'(a)x^a+\\&\quad+\sum_{N=1}^{\infty}x^{a+N}\Big\{A_N f'(a+N)+\frac{\partial A_N}{\partial a}f(a+N)-\\&\qquad\qquad-A_{N-1}g'(a+N-1)-\frac{\partial A_{N-1}}{\partial a}g(a+N-1)\Big\},\end{aligned}$$

with the convention $A_0 = 1$. But the coefficients A_N satisfy the recurrence-formula

$$A_N f(a+N) = A_{N-1}\, g(a+N-1),$$

and therefore, on differentiation,

$$A_N f'(a+N) + \frac{\partial A_N}{\partial a} f(a+N) = A_{N-1}\, g'(a+N-1) + \frac{\partial A_{N-1}}{\partial a} g(a+N-1).$$

Our preceding result therefore reduces simply to

$$\{f(\delta) - x\,g(\delta)\}\, S'(a) = f(a)x^a \log x + f'(a)x^a,$$

which is exactly (60). We can evidently deal in the same way with the results of operating with $f(\delta) - x\,g(\delta)$ on $S''(a)$, $S'''(a)$,... .

We have now only to consider the ordinary convergence of the series $S'(a)$, $S''(a)$. By (61)

$$\frac{\partial A_N}{\partial a} \Big/ A_N = -\sum_{r=1}^{N} \Big\{ \sum_{s=1}^{m} (a-a_s+r)^{-1} - \sum_{t=1}^{n} (a-b_t+r-1)^{-1} \Big\}$$

$$= -\sum_{r=1}^{N} \Big\{ \sum_{s=1}^{m} \frac{a_s - b_s - 1}{(a-a_s+r)(a-b_s+r-1)} \Big\}, \quad \text{if } m = n.$$

Thus, if $m = n$,

$$\frac{\partial A_N}{\partial a} \Big/ A_N \quad \text{is comparable with} \quad \sum^{N} r^{-2},$$

i.e. is bounded as $N \to \infty$, and the series $\sum (\partial A_N/\partial a) x^{a+N}$ converges with the series $\sum A_N x^{a+N}$, i.e. $S'(a)$ converges with $S(a)$, except for possible difficulties at $x = 0$ due to the factor $\log x$.

$$\text{If } m > n, \qquad \frac{\partial A_N}{\partial a} \Big/ A_N \quad \text{is comparable with} \quad \sum^{N} r^{-1},$$

i.e. with $\log N$. Thus $\sum (\partial A_N/\partial a) x_1^{a+N}$ converges more rapidly than $\sum A_N x^{a+N}$, if $|x_1| < |x|$. But, if $m > n$, $\sum A_N x^{a+N}$ converges in any finite interval, and therefore $\sum (\partial A_N/\partial a) x^{a+N}$ converges in any finite interval. Thus, if $m > n$, $S'(a)$ converges in any finite interval, possible difficulties at $x = 0$ excepted.

Thus, generally, $S'(a)$ has the same interval of convergence as $S(a)$. If this interval is finite and $S(a)$ converges conditionally but not absolutely at an end-point, the foregoing analysis is insufficient to decide the convergence of $S'(a)$ at this point, but this more delicate matter is best left for special discussion when it arises. By similar arguments we show that $S''(a)$, $S'''(a)$,... have the same intervals of convergence as $S(a)$. This completes the justification of the formal arguments of

§§ 11, 12 and at the same time completes the theory of the solution in power-series of the hypergeometric differential equation

$$f(\delta)y = x\,g(\delta)y.$$

By changing the dependent variable from x to x^c, where c is any constant, the theory can be applied with obvious modifications to the more general differential equation

$$f(\delta)y = x^c g(\delta)y.$$

14. A set of trigonometrical expansions

We illustrate the foregoing theory by applying it to obtain the expansions of $\sin m\theta$ and $\cos m\theta$, where m is any constant, in powers of $\sin\theta$ and of $\cos\theta$. Consider the function

$$y = \sin(m \sin^{-1} x+\alpha),$$

where α is a constant. Then

$$Dy = \frac{m\cos(m\sin^{-1}x+\alpha)}{\sqrt{(1-x^2)}},$$

i.e.
$$(1-x^2)(Dy)^2 = m^2(1-y^2).$$

A second differentiation gives

$$(1-x^2)\,D^2y - x\,Dy + m^2y = 0,$$

i.e.
$$\delta(\delta-1)y = x^2(\delta^2-m^2)y. \tag{70}$$

This differential equation has the two solutions in ascending powers of x

$$S_1(x) \equiv 1-\frac{m^2}{2!}x^2+\frac{m^2(m^2-2^2)}{4!}x^4-\frac{m^2(m^2-2^2)(m^2-4^2)}{6!}x^6+\ldots,$$

$$S_2(x) \equiv mx-\frac{m(m^2-1^2)}{3!}x^3+\frac{m(m^2-1^2)(m^2-3^2)}{5!}x^5-\ldots.$$

By chapter V (20) the general solution of the differential equation (70) is

$$y = A\,S_1(x)+B\,S_2(x),$$

where A, B are arbitrary constants. When $x = 0$ we have

$$y = \sin\alpha, \qquad S_1 = 1, \qquad S_2 = 0,$$

$$Dy = m\cos\alpha, \qquad DS_1 = 0, \qquad DS_2 = m.$$

This gives $A = \sin\alpha$, $B = \cos\alpha$ and so

$$\sin(m\sin^{-1}x+\alpha) = \sin\alpha\,S_1(x)+\cos\alpha\,S_2(x). \tag{71}$$

Put $x = \sin\theta$, $\alpha = 0$, $\frac{1}{2}\pi$, and we get respectively

$$\left.\begin{aligned} \sin m\theta &= S_2(\sin\theta) = m\sin\theta - \frac{m(m^2-1^2)}{3!}\sin^3\theta + \ldots \\ \cos m\theta &= S_1(\sin\theta) = 1 - \frac{m^2}{2!}\sin^2\theta + \frac{m^2(m^2-2^2)}{4!}\sin^4\theta - \ldots \end{aligned}\right\}. \tag{72}$$

Put $x = \cos\theta$, $\alpha = -\frac{1}{2}m\pi$, $\frac{1}{2}(1-m)\pi$, and we get respectively

$$\left.\begin{aligned} \sin m\theta &= \sin\tfrac{1}{2}m\pi\, S_1(\cos\theta) - \cos\tfrac{1}{2}m\pi\, S_2(\cos\theta) \\ \cos m\theta &= \cos\tfrac{1}{2}m\pi\, S_1(\cos\theta) + \sin\tfrac{1}{2}m\pi\, S_2(\cos\theta) \end{aligned}\right\}. \tag{73}$$

We can, of course, simplify (73), when m is an integer, if we distinguish the cases of m odd and m even.

We remark that the series S_1 terminates, if and only if m is an even integer; the series S_2 terminates, if and only if m is an odd integer. These are merely the known facts that $\cos 2m\theta$, $\sin(2m+1)\theta$ are expressible as polynomials in $\sin\theta$, and $\cos 2m\theta$, $\cos(2m+1)\theta$ as polynomials in $\cos\theta$.

If the series $S_1(x)$, $S_2(x)$ do not terminate, we find from (59) that they converge in the closed interval $|x| \leqslant 1$. For, if in (70) we put $x' \equiv x^2$, we get

$$\delta'(\delta'-\tfrac{1}{2})y = x'(\delta'^2 - \tfrac{1}{4}m^2)y, \tag{74}$$

in which $p = 1 = q$, $2+p'/p-q'/q = \frac{3}{2} > 1$, so that there is convergence at the end-point $x' = 1$, i.e. at $x = \pm 1$.

Thus the series (72), (73) converge for all real values of θ: in particular, if in (72) we put $\theta = \frac{1}{2}\pi$, we get the series

$$\left.\begin{aligned} \sin\tfrac{1}{2}m\pi &= m - \frac{m(m^2-1^2)}{3!} + \frac{m(m^2-1^2)(m^2-3^2)}{5!} - \ldots \\ \cos\tfrac{1}{2}m\pi &= 1 - \frac{m^2}{2!} + \frac{m^2(m^2-2^2)}{4!} - \ldots \end{aligned}\right\}. \tag{75}$$

In comparison with these it is interesting to note that the expressions for $\sin\frac{1}{2}m\pi$ and $\cos\frac{1}{2}m\pi$ as infinite products can also be written in the form of infinite series,† namely

$$\left.\begin{aligned} \cos\tfrac{1}{2}m\pi &= 1 - \frac{m^2}{1^2} + \frac{m^2(m^2-1^2)}{1^2.3^2} - \frac{m^2(m^2-1^2)(m^2-3^2)}{1^2.3^2.5^2} + \ldots \\ \sin\tfrac{1}{2}m\pi &= \tfrac{1}{2}m\pi\left[1 - \frac{m^2}{2^2} + \frac{m^2(m^2-2^2)}{2^2.4^2} - \frac{m^2(m^2-2^2)(m^2-4^2)}{2^2.4^2.6^2} + \ldots\right] \end{aligned}\right\}. \tag{75'}$$

† These series (75′) are obtained from the series (72) above by integrating in θ from 0 to $\frac{1}{2}\pi$. This gives a simple method of establishing these infinite products.

15. The associated descending series

For the solution of (70) in *descending* powers we obtain, after reduction, the two series

$$\left.\begin{aligned}
&T_1(x)\\
&\equiv x^m-\frac{m}{4}x^{m-2}+\frac{m(m-3)}{4.8}x^{m-4}-\frac{m(m-4)(m-5)}{4.8.12}x^{m-6}+\dots\\
&T_2(x)\\
&\equiv x^{-m}+\frac{m}{4}x^{-(m+2)}+\frac{m(m+3)}{4.8}x^{-(m+4)}+\frac{m(m+4)(m+5)}{4.8.12}x^{-(m+6)}+\dots
\end{aligned}\right\} \quad (76)$$

which differ only in the sign of m. Since m enters the differential equation (70) only through its square, we can sufficiently take m to be positive.

In particular, if m is a positive integer, then, in (74), the two roots $\pm\frac{1}{2}m$ of the 'descending' indicial equation differ by an integer and we should expect logarithmic terms. Happily, these two roots $\pm\frac{1}{2}m$ are separated (according as m is even or odd) by one or other of the roots $0, \frac{1}{2}$ of the 'ascending' indicial equation, and, in consequence, the series corresponding to $\frac{1}{2}m$ terminates. Now, if, in $T_1(x)$, we make m a positive integer, we find that the series does not terminate, but only that certain terms vanish in the middle. There remains a finite series followed by an infinite series. The infinite series is, naturally, some constant multiple of $T_2(x)$, which remains a solution whether m be an integer or not; and the finite series is, of course, the finite series we have just considered. If we write $T_0(x)$ for this finite series, we obtain, in actual fact, the identity

$$T_0(x) = T_1(x)+4^{-m}T_2(x) \quad (m \text{ a positive integer}). \tag{77}$$

Reference to the differential equation in the form (74) shows that the descending series converge only for $|x| \geqslant 1$. They do not therefore lead to expansions in $\sin\theta$ or $\cos\theta$, unless they terminate. This happens, only if m is an integer, and then only for the one series $T_0(x)$. Clearly $T_0(x)$ represents the terminated series $S_1(x)$ or $S_2(x)$, according as m is even or odd, written in the reverse order. The $(n+1)$th term in $T_1(x)$ is

$$(-)^n\frac{m(m-n-1)(m-n-2)\dots(m-2n+1)}{4.8.12\dots4n}x^{m-2n}.$$

Thus, according as $m = 2n$ or $2n+1$, the last term in $T_0(x)$ is

$$(-)^{\frac{1}{2}m}/2^{m-1} \quad \text{or} \quad (-)^{\frac{1}{2}(m-1)}mx/2^{m-1}.$$

Hence, by comparison of lowest terms, we have

$$\left.\begin{aligned} 2^{m-1}T_0(x) &= (-)^{\frac{1}{2}m}\,S_1(x) \qquad (m \text{ even}) \\ &= (-)^{\frac{1}{2}(m-1)}\,S_2(x) \quad (m \text{ odd}) \end{aligned}\right\}. \tag{78}$$

If we put in succession $x = \sin\theta$, $x = \cos\theta$ and refer to (72) and (73), we find that (78) reduces to the well-known trigonometrical formulae

$$\begin{aligned} &(2\sin\theta)^m - m(2\sin\theta)^{m-2} + \frac{m(m-3)}{2!}(2\sin\theta)^{m-4} - \ldots \\ &\quad \left.\begin{aligned} = 2^mT_0(\sin\theta) &= (-)^{\frac{1}{2}m}2\cos m\theta \qquad (m \text{ even}) \\ &= (-)^{\frac{1}{2}(m-1)}\sin m\theta \quad (m \text{ odd}) \end{aligned}\right\}, \end{aligned} \tag{79}$$

and

$$\begin{aligned} &(2\cos\theta)^m - m(2\cos\theta)^{m-2} + \frac{m(m-3)}{2!}(2\cos\theta)^{m-4} - \ldots \\ &\qquad = 2^mT_0(\cos\theta) = 2\cos m\theta \qquad (m \text{ an integer}). \end{aligned} \tag{80}$$

To obtain an interpretation for the infinite series $T_1(x)$, $T_2(x)$ when $|x| > 1$, consider the function

$$z = \{x+\surd(x^2-1)\}^m, \tag{81}$$

which is real when $x > 1$, or, if m is an integer, when $|x| \geqslant 1$. Differentiation of $\log z$ gives

$$\surd(x^2-1)\,Dz = mz.$$

A second differentiation gives

$$\surd(x^2-1)\,D^2z + \frac{x}{\surd(x^2-1)}\,Dz = m\,Dz,$$

and so
$$(x^2-1)\,D^2z + x\,Dz = m^2z,$$

which is equation (70).

Since m enters the differential equation only through its square, this gives the two solutions

$$\{x+\surd(x^2-1)\}^{\pm m},$$

which are therefore linear functions of the two other solutions $T_1(x)$, $T_2(x)$. When $x \to \infty$, we have, to a first approximation,

$$T_1(x) \sim x^m, \qquad T_2(x) \sim x^{-m}, \qquad \{x+|\surd(x^2-1)|\}^{\pm m} \sim (2x)^{\pm m}.$$

Thus we can evidently make the identifications

$$\{x+|\surd(x^2-1)|\}^{-m} = \{x-|\surd(x^2-1)|\}^m \ = 2^{-m}T_2(x), \tag{82}$$

$$\{x+|\surd(x^2-1)|\}^m \ = \{x-|\surd(x^2-1)|\}^{-m} = 2^mT_1(x)+AT_2(x), \tag{83}$$

where A is some constant.

If m is a positive integer,

$$\{x-\surd(x^2-1)\}^m+\{x+\surd(x^2-1)\}^m$$

is clearly expansible as a polynomial in x, and is thus some constant multiple of $T_0(x)$. Using (77) we make the identification

$$\{x-\sqrt{(x^2-1)}\}^m+\{x+\sqrt{(x^2-1)}\}^m = 2^mT_1(x)+2^{-m}T_2(x) = 2^mT_0(x), \quad (84)$$

and A is zero.

If m is not an integer, we can still see that A must be zero, for

$$\{x+\sqrt{(x^2-1)}\}^m = x^m\{1+\sqrt{(1-x^{-2})}\}^m$$

and so, by an appeal to Taylor's theorem, is expansible in a series of powers x^{m-2n}, where n is a positive integer. But, on the right of (83), $T_2(x)$ will contribute powers x^{-m-2n}, and these are not of the form x^{m-2n} unless m is an integer. We must therefore also have $A = 0$ when m is not an integer, and we can accordingly supersede (82), (83) by the single equation

$$(2x)^m-m(2x)^{m-2}+\frac{m(m-3)}{2!}(2x)^{m-4}-\frac{m(m-4)(m-5)}{3!}(2x)^{m-6}+\ldots \quad \text{to } \infty$$

$$= \{x+\sqrt{(x^2-1)}\}^m = \{x-\sqrt{(x^2-1)}\}^{-m} = 2^mT_1(x), \quad (85)$$

which is true for all values of m.

It should be remarked that (85) can also be obtained as a Lagrange's expansion of $(xy)^m$ in ascending powers of $(4x^2)^{-1}$, where y is defined implicitly by the (quadratic) equation

$$y = 1-(4x^2)^{-1}y^{-1}$$

and is the root that converges to 1 as $(4x^2)^{-1}$ converges to 0.

To obtain from (85) some sort of analogy with the trigonometrical series (72), (73), write $x = \cosh\theta$ in (85) and we have

$$e^{m\theta} = (2\cosh\theta)^m-m(2\cosh\theta)^{m-2}+\frac{m(m-3)}{2!}(2\cosh\theta)^{m-4}-$$

$$-\frac{m(m-4)(m-5)}{3!}(2\cosh\theta)^{m-6}+\ldots \quad \text{to } \infty \quad (\theta > 0). \quad (86)$$

Finally, since $T_1(x)$ converges up to $|x| = 1$, we may put $x = 1$ in (85) and obtain the arithmetical identity

$$1 = 2^m-m2^{m-2}+\frac{m(m-3)}{2!}2^{m-4}-\frac{m(m-4)(m-5)}{3!}2^{m-6}+\ldots \text{ to } \infty. \quad (87)$$

We can deduce new expansions from the foregoing by differentiation. Thus differentiation with respect to x gives in (71) and (85) the expansions of

$$\frac{\sin(m\sin^{-1}x+\alpha)}{\sqrt{(1-x^2)}} \quad \text{and} \quad \frac{\{x+\sqrt{(x^2-1)}\}^m}{\sqrt{(x^2-1)}}.$$

These two functions are both solutions of the differential equation

$$\delta(\delta-1)y = x^2\{(\delta+1)^2-m^2\}y$$

and their expansions could, of course, be deduced directly from this fact. The substitutions $x = \sin\theta$ and $x = \cos\theta$ then give the expansions of

$$\frac{\sin m\theta}{\cos\theta}, \quad \frac{\cos m\theta}{\cos\theta} \quad \text{in powers of } \sin\theta,$$

and of

$$\frac{\sin m\theta}{\sin\theta}, \quad \frac{\cos m\theta}{\sin\theta} \quad \text{in powers of } \cos\theta.$$

These are, of course, the expansions obtained by differentiating (72) and (73) with respect to θ Similarly, differentiation of (86) gives the expansion of $e^{m\theta}/\sinh\theta$ in powers of $\cosh\theta$. It will be found that these differentiated series do not converge at the end-points of the interval of convergence, and we do not therefore obtain arithmetical series analogous to (75) and (87).

16. Postscriptum.

We can deduce from equation (20) of § 4 above a number of interesting trigonometrical expansions. In that equation write $\alpha = \pi+x$, $n = 0$, and we have

$$\sum_{n=1}^{\infty} (-)^{n-1}\frac{\sin nx}{n} = \tfrac{1}{2}x \qquad (|x| < \pi). \tag{88}$$

The coefficients in the original series (20) converge monotonically to zero, and therefore, as pointed out in chapter I § 13, that series converges uniformly over any closed interval that contains none of the points $0, \pm 2\pi, \pm 4\pi,\dots$. The series (88) thus converges uniformly over the closed interval $|x| \leqslant \pi-\epsilon$, for any small positive ϵ.

Write

$$C_2(x) \equiv \sum_{n=1}^{\infty} (-)^{n-1}\frac{\cos nx}{n^2}, \qquad S_3(x) \equiv \sum_{n=1}^{\infty} (-)^{n-1}\frac{\sin nx}{n^3}.$$

Then term-by-term differentiation gives

$$\frac{d}{dx}C_2(x) = -\sum_{n=1}^{\infty} (-)^{n-1}\frac{\sin nx}{n}.$$

This differentiation is permissible in the interval $|x| \leqslant \pi-\epsilon$, since the differentiated series converges uniformly over that interval. Therefore, by (88), in the interval $|x| < \pi$,

$$\frac{d}{dx}C_2(x) = -\tfrac{1}{2}x,$$

and so, in that interval,

$$C_2(x) = c-\tfrac{1}{4}x^2, \tag{89}$$

where c is some constant. Now the series $C_2(x)$ is comparable with the convergent series of positive constants $\sum n^{-2}$ and so converges uniformly over any finite interval. It is therefore continuous, in particular at $x = \pm\pi$, and, by taking the limits $x \to \pi-$, $-\pi+$, we can extend (89) to the complete interval $|x| \leqslant \pi$.

Again,
$$\frac{d}{dx} S_3(x) = C_2(x),$$
the term-by-term differentiation being everywhere valid, since the differentiated series $C_2(x)$, as we have said, converges uniformly over any finite interval. Thus
$$S_3(x) = cx - \tfrac{1}{12}x^3 \qquad (|x| \leqslant \pi), \tag{90}$$
the additive constant being zero, since $S_3(0) = 0$. Since also $S_3(\pi) = 0$, we get $c = \frac{1}{12}\pi^2$, and so from (89), (90)
$$\left.\begin{aligned} \sum_{n=1}^{\infty} (-)^{n-1} \frac{\cos nx}{n^2} &= \frac{\pi^2 - 3x^2}{12} \\ \sum_{n=1}^{\infty} (-)^{n-1} \frac{\sin nx}{n^3} &= \frac{x(\pi^2 - x^2)}{12} \end{aligned}\right\} \quad (|x| \leqslant \pi). \tag{91}$$
In the first of these equations put successively $x = \pi$, 0, and we get
$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}, \qquad \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^2} = \frac{\pi^2}{12}. \tag{92}$$
From (91) we obtain by a similar series of steps
$$\left.\begin{aligned} \sum_{n=1}^{\infty} (-)^{n-1} \frac{\cos nx}{n^4} &= \frac{15x^4 - 30x^2\pi^2 + 7\pi^4}{720} \\ \sum_{n=1}^{\infty} (-)^{n-1} \frac{\sin nx}{n^5} &= \frac{x(x^2 - \pi^2)(3x^2 - 7\pi^2)}{720} \end{aligned}\right\} \quad (|x| \leqslant \pi), \tag{93}$$
whence
$$\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}, \qquad \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^4} = \frac{7\pi^4}{720}, \tag{94}$$
and so on.

Again, if p is not an integer, write
$$f(x) = \sum_{n=1}^{\infty} (-)^{n-1} \frac{\sin nx}{n(n^2 - p^2)}. \tag{95}$$
Then
$$f'(x) = \sum_{n=1}^{\infty} (-)^{n-1} \frac{\cos nx}{n^2 - p^2}, \tag{96}$$
the term-by-term differentiation being everywhere valid, since (96) is comparable with the convergent series of positive constants $\sum n^{-2}$ and

therefore converges uniformly over any finite interval. A second differentiation gives

$$(D^2+p^2)f(x) = -\sum_{n=1}^{\infty}(-)^{n-1}\frac{\sin nx}{n}. \tag{97}$$

Such a differentiation, as we have seen above, is valid in the open interval $)-\pi,\pi($, and so, in that interval,

$$(D^2+p^2)f(x) = -\tfrac{1}{2}x,$$

i.e.

$$(D^2+p^2)\{f(x)-x/2p^2\} = 0. \tag{98}$$

Now $y = \sin px$, $y = \cos px$ are two linearly independent solutions of the differential equation

$$(D^2+p^2)y = 0,$$

and so, by chapter V (20) the general solution is

$$y = A\sin px+B\cos px,$$

where A, B are independent of x. Thus, from (98),

$$f(x) = A\sin px+B\cos px-x/2p^2. \tag{99}$$

This holds throughout the open interval $)-\pi,\pi($, but by an argument used above we can extend its validity to the end-points $\pm\pi$, since both sides of the equation are continuous at these points. Since $f(x)$ vanishes at $x = 0$, $\pm\pi$, we have, from (99),

$$A = \frac{\pi}{2p^2}\operatorname{cosec} p\pi, \qquad B = 0,$$

and accordingly from (96), (95)

$$\left.\begin{aligned}\sum_{n=1}^{\infty}(-)^{n-1}\frac{\cos nx}{n^2-p^2} &= \frac{\pi p\operatorname{cosec} p\pi\cos px-1}{2p^2}\\ \sum_{n=1}^{\infty}(-)^{n-1}\frac{\sin nx}{n(n^2-p^2)} &= \frac{\pi\operatorname{cosec} p\pi\sin px-x}{2p^2}\end{aligned}\right\} \quad (|x| \leqslant \pi). \tag{100}$$

If in the first of these equations we put successively $x = \pi$, 0, we have after slight rearrangement

$$\pi\cot p\pi = \frac{1}{p}+\sum_{n=1}^{\infty}\frac{2p}{p^2-n^2}, \qquad \pi\operatorname{cosec} p\pi = \frac{1}{p}+\sum_{n=1}^{\infty}(-)^n\frac{2p}{p^2-n^2}. \tag{101}$$

The first of these equations (101) may be written

$$\begin{aligned}\frac{d}{dp}\log\frac{\sin p\pi}{p} &= \sum_{n=1}^{\infty}\frac{d}{dp}\log\left(1-\frac{p^2}{n^2}\right)\\ &= \frac{d}{dp}\sum_{n=1}^{\infty}\log\left(1-\frac{p^2}{n^2}\right),\end{aligned}$$

since term-by-term differentiation is permissible, the differentiated series being comparable, for convergence, with the convergent series $\sum n^{-2}$. Thus

$$\log\frac{\sin p\pi}{p} = A+\sum_{n=1}^{\infty}\log\left(1-\frac{p^2}{n^2}\right),$$

where A is independent of p. The series on the right is again comparable with $\sum n^{-2}$ and is therefore continuous in p (integer values of p alone excepted). In particular, we may take the limit $p\to 0$ on both sides of the equation, which gives $A=\log\pi$, and so

$$\sin p\pi = p\pi\prod_{n=1}^{\infty}\left(1-\frac{p^2}{n^2}\right).$$

We write this in more usual form, of course, as

$$\sin\theta = \theta\prod_{n=1}^{\infty}\left(1-\frac{\theta^2}{n^2\pi^2}\right).$$

The corresponding infinite product for $\cos\theta$

$$\cos\theta = \prod_{n=1}^{\infty}\left(1-\frac{\theta^2}{(n+\frac{1}{2})^2}\right)$$

may be deduced from that for $\sin\theta$ by writing $\pi+\theta$ for θ therein or obtained, as above, from equations (101) by noting that

$$-\tfrac{1}{2}\pi\tan\tfrac{1}{2}p\pi = \tfrac{1}{2}\pi(\cot p\pi-\operatorname{cosec}p\pi) = \sum_{n=1}^{\infty}\frac{2p}{p^2-(2n+1)^2}.$$

We may differentiate equations (100) and (101) with respect to p, for the resulting series are comparable, for convergence, with the series $\sum n^{-4}$. It is convenient to write the equations in the modified form

$$\frac{x}{p}+\sum_{n=-\infty}^{\infty}{}'(-)^n\frac{\sin nx}{n(p+n)} = \frac{\pi\operatorname{cosec}p\pi\sin px}{p},$$

$$\sum_{n=-\infty}^{\infty}(-)^n\frac{\cos nx}{p+n} = \pi\operatorname{cosec}p\pi\cos px,$$

$$\sum_{n=-\infty}^{\infty}\frac{1}{p+n} = \pi\cot p\pi, \qquad \sum_{n=-\infty}^{\infty}\frac{(-1)^n}{p+n} = \pi\operatorname{cosec}p\pi.$$

Then differentiation in p gives

$$\frac{x}{p^2}+\sum_{n=-\infty}^{\infty}{}'(-)^n\frac{\sin nx}{n(p+n)^2} = \frac{\pi\operatorname{cosec}p\pi}{p^2}\{(1+p\pi\cot p\pi)\sin px-px\cos px\},$$

$$\sum_{n=-\infty}^{\infty}(-)^n\frac{\cos nx}{(p+n)^2} = \pi\operatorname{cosec}p\pi(x\sin px+\pi\cot p\pi\cos px),$$

$$\sum_{n=-\infty}^{\infty} \frac{1}{(p+n)^2} = \pi^2 \operatorname{cosec}^2 p\pi, \qquad \sum_{n=-\infty}^{\infty} \frac{(-1)^n}{(p+n)^2} = \pi^2 \operatorname{cosec} p\pi \cot p\pi,$$

and we can proceed in this way by further differentiations to results of increasing complexity.

In the above summations $\sum'$ omits the term $n = 0$, and, in the first set,

$$\sum_{n=-\infty}^{\infty} \text{ means } \lim_{N\to\infty} \sum_{n=-N}^{N},$$

while, in the second set, $\pm\infty$ can be independent, since the various series are convergent.

It should be remarked, in conclusion, that the sum-functions of the trigonometric series considered above are analytic only in restricted intervals. They are differentiable, with derivatives everywhere continuous, up to, but only up to, a certain order of differentiation. In each of the intervals in question, the sum-function is identical with a certain polynomial, but this polynomial is different in successive intervals. Thus the sum-function and the polynomial are distinct functions that have the same analytical expression in a certain interval, a fact which emphasizes very sharply the difference between a 'function' and its 'analytical expression', already noted at the beginning of this book.

WORKED EXAMPLE

If m is an integer greater than 2, solve the algebraic trinomial equation

$$xy^m - y + 1 = 0$$

by means of hypergeometric series $y = y(x)$.

Writing the given equation as an equation implicit in y

$$y = 1 + xy^m \tag{1}$$

and applying the theory of Lagrange's series (§ 2 above), we get at once the expansion

$$\begin{aligned} y &= 1 + \sum_{n=1}^{\infty} \frac{x^n}{n!} \left[\frac{d^{n-1}(y^{mn})}{dy^{n-1}}\right]_{y=1} \\ &= \sum_{n=0}^{\infty} \frac{(mn)!\, x^n}{n!\{(m-1)n+1\}!} \\ &= \sum_{n=0}^{\infty} A_n x^n, \quad \text{say,} \end{aligned}$$

where

$$\frac{A_{n+1}}{A_n} = \frac{(mn+1)(mn+2)\ldots(mn+m)}{(n+1)\{(m-1)n+2\}\{(m-1)n+3\}\ldots\{(m-1)n+m\}}$$
$$= \frac{m(mn+1)(mn+2)\ldots(mn+m-1)}{\{(m-1)(n+1)+3-m\}\{(m-1)(n+1)+4-m\}\ldots\{(m-1)(n+1)+1\}}.$$

Thus $\sum A_n x^n$ is a hypergeometric series and a solution of the hypergeometric differential equation

$$\{(m-1)\delta+3-m\}\{(m-1)\delta+4-m\}\ldots(m-1)\delta\{(m-1)\delta+1\}y$$
$$= mx(m\delta+1)(m\delta+2)\ldots(m\delta+m-1)y, \quad (2)$$

which we may write compactly as

$$\prod_{s=1}^{m-1}\left(\delta-1+\frac{s+1}{m-1}\right)y = \frac{m^m x}{(m-1)^{m-1}}\prod_{s=1}^{m-1}\left(\delta+\frac{s}{m}\right)y. \quad (3)$$

We may, of course, obtain this differential equation direct from the algebraic equation. For operation with the operator $y^{1-m}\delta$ on (1) gives

$$y^{1-m}\,\delta y = x(m\delta+1)y. \quad (4)$$

But, again from (1), $\quad y^{1-m} = xy/(y-1)$,

which gives in (4) $\quad y\,\delta y = (y-1)(m\delta+1)y$,

i.e. $\quad y\{(m-1)\delta+1\}y = (m\delta+1)y$.

If we operate on this with the operator

$$y^{s-2}\{(s-1)(\delta y)+y\delta\}$$

and reduce, we get

$$\{(m-1)\delta+s+1\}y^s\,\delta y = (m\delta+s)y^{s-1}\,\delta y, \quad (5)$$

which we use as a recurrence-formula, s being a parameter at our disposal. Then

$$mx\prod_{s=1}^{m-1}(m\delta+s)y = m\prod_{s=2}^{m-1}(m\delta-m+s).x(m\delta+1)y$$

$$= m\prod_{s=3}^{m-1}(m\delta-m+s).(m\delta-m+2)y^{1-m}\,\delta y, \quad \text{by (4)},$$

$$= m\prod_{s=3}^{m-1}(m\delta-m+s).\{(m-1)\delta+3-m\}y^{2-m}\,\delta y, \quad \text{by (5) with } s=2-m,$$

$$= m\{(m-1)\delta+3-m\}\prod_{s=4}^{m-1}(m\delta-m+s).(m\delta-m+3)y^{2-m}\,\delta y$$

$$= m\{(m-1)\delta+3-m\}\prod_{s=4}^{m-1}(m\delta-m+s).\{(m-1)\delta+4-m\}y^{3-m}\,\delta y,$$

by (5) with $s=3-m$, and so on.

Thus, at length, by taking further $s = 4-m, 5-m, \ldots, -1$ in (5), we obtain

$$mx\prod_{s=1}^{m-1}(m\delta+s)y = \prod_{s=1}^{m-3}\{(m-1)\delta-m+s+2\}.m(m-1)\,\delta y^{-1}\,\delta y.$$

But $\quad m(m-1)\delta y^{-1}\,\delta y = (m-1)\{(m-1)\delta+1\}\,\delta y$, by (5) with $s=0$,

and so finally we have the required formula

$$mx\prod_{s=1}^{m-1}(m\delta+s)y = \prod_{s=1}^{m-1}\{(m-1)\delta-m+s+2\}y.$$

Examination of the indicial equations of this differential equation shows that there are ascending series with leading exponents

$$r = -\frac{1}{m-1},\quad 0,\quad \frac{1}{m-1},\quad \ldots,\quad \frac{m-3}{m-1}$$

and descending series with leading exponents

$$r = -\frac{1}{m},\quad -\frac{2}{m},\quad \ldots,\quad -\frac{m-1}{m}.$$

We may write the ascending series $S_r(x)$ in the form

$$S_p(x) = x^r \sum_{n=0}^{\infty} \left\{\frac{m^m x}{(m-1)^{m-1}}\right\}^n \prod_{s=1}^{m-1} \prod_{t=0}^{n-1} \frac{r+t+s/m}{r+t+(s+1)/(m-1)}$$

$$\left(p = 1, 2, \ldots, m-1;\ r = -\frac{1}{m-1}, 0, \ldots, \frac{m-3}{m-1};\ p \equiv (m-1)r+2\right) \quad (6)$$

and the descending series in the form

$$T_p(x) = x^r \sum_{n=0}^{\infty} \left\{\frac{(m-1)^{m-1}}{m^m x}\right\}^n \prod_{s=1}^{m-1} \prod_{t=0}^{n-1} \frac{r-t-1+(s+1)/(m-1)}{r-t-1+s/m}$$

$$\left(p = 1, 2, \ldots, m-1;\ r = -\frac{1}{m}, -\frac{2}{m}, \ldots, -\frac{m-1}{m};\ p = -mr\right), \quad (7)$$

being careful in each case to retain unity as the leading coefficient.

For brevity write $\qquad M = \dfrac{(m-1)^{m-1}}{m^m}.$

Then the ascending series S_r converge in the interval $|x| < M$; the descending series T_r converge outside this interval. If $m > 2$, the tests of (59) above show that the ascending series and the descending series both converge at the end-points of this interval.

To determine the numbers of real and of imaginary roots of the equation write

$$f(y) \equiv y^m - x^{-1}y + x^{-1}.$$

Then

$$f'(y) = my^{m-1} - x^{-1},$$
$$f''(y) = m(m-1)y^{m-2}.$$

If m is even, there is a single stationary value y_0 given by $f'(y_0) = 0$, and $f''(y)$ is invariably positive. The stationary value is therefore a minimum and the equation has two real roots or no real roots according as $f(y_0) < 0$ or > 0. But

$$f(y_0) = x^{-1}\{1-(M/x)^{1/(m-1)}\}.$$

Thus the equation has two real roots or no real root according as $M > x$ or $< x$. In the limiting case $M = x$ the two real roots become coincident.

If m is odd and $x < 0$, then $f'(y)$ is invariably positive, $f(y)$ is monotonic, and the equation has a single real root. If m is odd and $x > 0$, then there are two stationary values $\pm y_0$, say. Since $f''(y)$ has now the sign of y, the upper sign gives a minimum, the lower sign a maximum. Again

$$f(\pm y_0) = x^{-1}\{1 \mp (M/x)^{1/(m-1)}\}.$$

Thus f is always positive at the maximum, at the minimum it is positive or negative according as $x > M$ or $< M$.

We may sum up our results as follows:

m even,	$x > M$:	no real root,
	$x < M$:	two real roots;
m odd,	$x > M$:	one real root,
	$0 < x < M$:	three real roots,
	$x < 0$:	one real root.

(8)

Since the differential equation has been deduced direct from the algebraic

equation, it is clear that every root of the algebraic equation (regarded as a function of x) is also a solution of the differential equation. Now the algebraic equation has m distinct roots (if $x \neq M$): the differential equation has only $m-1$ linearly independent solutions. This apparent discrepancy is reconciled when we observe that, if $m > 2$, the sum of the roots of the algebraic equation is zero, i.e. these m roots are linearly connected and therefore count as only $m-1$ linearly independent solutions of the differential equation.

We have proved, then, that the roots of the algebraic equation are expressible in the form†

$$y = a_1 S_1 + a_2 S_2 + \ldots + a_{m-1} S_{m-1} \qquad (|x| < M)$$

or else

$$y = b_1 T_1 + b_2 T_2 + \ldots + b_{m-1} T_{m-1} \qquad (|x| > M),$$

where a_r, b_r are constants. To determine these constants we could substitute into the algebraic equation and consider dominant terms as $x \to 0$ and $x \to \infty$ respectively. But it is simpler to expand the several roots direct in Lagrange's series.‡

We began our discussion with the Lagrange's expansion of that root of the algebraic equation that converges to 1 as x converges to 0. This turned out to be exactly the hypergeometric series that we have called S_2. To obtain the Lagrange's expansions of the other $m-1$ roots in the region $|x| < M$ (which contains the point $x = 0$) write

$$Y = xy^{m-1},$$

so that

$$y = \omega^{-1}x^{-1/(m-1)}Y^{1/(m-1)}, \quad \text{where } \omega^{m-1} = 1,$$

and the algebraic equation becomes

$$Y = 1 - \omega x^{1/(m-1)}Y^{-1/(m-1)}.$$

The corresponding Lagrange's series for Y gives for y the expansion

$$y = \omega^{-1}x^{-1/(m-1)}\left\{1 + \frac{1}{m-1}\sum_{n=1}^{\infty}\frac{(-\omega x^{1/(m-1)})^n}{n!}\left[\frac{d^{n-1}}{dY^{n-1}}Y^{-(m+n-2)/(m-1)}\right]_{Y=1}\right\}$$

$$= \omega^{-1}x^{-1/(m-1)} -$$

$$-\sum_{n=1}^{\infty}\frac{(m+n-2)(2m+n-3)\ldots\{(n-2)m+1\}(n-1)m}{(m-1)^n n!}(\omega x^{1/(m-1)})^{n-1}.$$

By identifying the first $m-1$ terms of this expansion with leading coefficients in the S_r we have the result

$$y = \omega^{-1}S_1(x) - \sum_{n=1}^{m-2}\frac{(m+n-2)(2m+n-3)\ldots\{(n-2)m+1\}(n-1)m}{(m-1)^n n!}\omega^{n-1}S_{n+1}(x), \quad (9)$$

valid when $|x| \leqslant M$, where $\omega^{m-1} = 1$. This, together with $y = S_2$, gives the m roots in the region $|x| \leqslant M$. Of these S_2 is always real, while (9) gives one real root, if m is even: two real roots, if m is odd and x positive: and no real root, if m is odd and x negative, for then S_1, for instance, has the imaginary factor $x^{-1/(m-1)}$. This agrees with (8) above.

† We must not, of course, fall into the error of supposing that the S_r or the T_r that we have taken as the fundamental solutions of the differential equation are necessarily also the actual roots of the algebraic equation.

‡ This will not be strictly legitimate, since it presumes the validity of Lagrange's expansion for a complex argument.

In the region $|x| > M$ write

$$Y \equiv -xy^m,$$

so that $$y = \omega_1 x^{-1/m} Y^{1/m}, \quad \text{where} \quad \omega_1^m = -1,$$

and the algebraic equation becomes

$$Y = 1 - \omega_1 x^{-1/m} Y^{1/m}.$$

The corresponding Lagrange's series for Y gives for y the expansion

$$y = \omega_1 x^{-1/m}\left\{1 + \frac{1}{m}\sum_{n=1}^{\infty}\frac{(-\omega_1 x^{-1/m})^n}{n!}\left[\frac{d^{n-1}}{dY^{n-1}} Y^{-(m-n-1)/m}\right]_{Y=1}\right\}$$

$$= \omega_1 x^{-1/m} - \sum_{n=1}^{\infty}\frac{(m-n-1)(2m-n-1)\ldots\{(n-1)m-n-1\}}{m^n n!}(\omega_1 x^{-1/m})^{n+1}$$

from which we derive the result

$$y = \omega_1 T_1(x) - \sum_{n=1}^{\infty}\frac{(m-n-1)(2m-n-1)\ldots\{(n-1)m-n-1\}}{m^n n!}\omega_1^{n+1} T_{n+1}(x), \quad (10)$$

valid when $|x| > M$, where $\omega_1^m = -1$. This gives the m roots in the region $|x| > M$. None or one of them will be real according as m is even or odd, which is in accordance with (8) above.

This completes the theory of the solution of the trinomial equation in terms of hypergeometric series.

EXAMPLES XIV

1. Expand in ascending powers of x the functions

$$\text{(i)}\ \log\frac{1+x\sqrt{2}+x^2}{1-x\sqrt{2}+x^2} + 2\tan^{-1}\frac{x\sqrt{2}}{1-x^2},$$

$$\text{(ii)}\ \log\frac{1+x+x^2}{1-2x+x^2} + 2\sqrt{3}\tan^{-1}\frac{2x+1}{\sqrt{3}},$$

$$\text{(iii)}\ \log\frac{x^2-1+\sqrt{(1+x^4)}}{x^2}.$$

2. Obtain in ascending powers of x the approximate expansions of

$\tan^2 x$ (to x^{10}), $\quad \cot^2 x$ (to x^8),

$\tan x \sec x$ (to x^7), $\quad \cot x \operatorname{cosec} x$ (to x^4),

$\sec^3 x$ (to x^6), $\quad \operatorname{cosec}^3 x$ (to x^3).

3. If r be a positive integer and

$$\frac{(1+x)^r \log(1+x)}{x} = a_0 + a_1 x + a_2 x^2 + \ldots \quad \text{to } \infty,$$

show that

$$\text{(i)}\ a_n > 0 \qquad (n \leqslant r),$$

$$\text{(ii)}\ a_n = (-)^{n-r}\frac{(n-r)!\,r!}{(n+1)!} \qquad (n \geqslant r).$$

4. If $m > 0$, $1 > x > 0$, and

$$f(x) \equiv \frac{1}{m}\frac{1}{1+x} + \frac{1!}{m(m+1)}\frac{x}{(1+x)^2} + \frac{2!}{m(m+1)(m+2)}\frac{x^2}{(1+x)^3} + \ldots \quad \text{to } \infty,$$

show that $$xf'(x) + mf(x) = \frac{1}{1+x},$$

and obtain the expansion

$$f(x) = \frac{1}{m} - \frac{x}{m+1} + \frac{x^2}{m+2} - \dots .$$

5. (i) If

$$(1 - a_1 x + a_2 x^2 - a_3 x^3 + \dots)^{-p} = 1 + b_1 x + \frac{b_2 x^2}{2!} + \frac{b_3 x^3}{3!} + \dots,$$

show that

$$b_n = p \begin{vmatrix} a_1 & 1 & 0 & 0 & . \\ 2a_2 & (p+1)a_1 & 2 & 0 & . \\ 3a_3 & (2p+1)a_2 & (p+2)a_1 & 3 & . \\ . & . & . & . & . \\ na_n & \{(n-1)p+1\}a_{n-1} & \{(n-2)p+2\}a_{n-2} & \{(n-3)p+3\}a_{n-3} & . \end{vmatrix} .$$

(ii) If $\log(1 - a_1 x + a_2 x^2 - a_3 x^3 + \dots) = -b_1 x - \frac{1}{2} b_2 x^2 - \frac{1}{3} b_3 x^3 - \dots$, show that

$$b_n = \begin{vmatrix} a_1 & 1 & 0 & 0 & . \\ 2a_2 & a_1 & 1 & 0 & . \\ 3a_3 & a_2 & a_1 & 1 & . \\ . & . & . & . & . \\ na_n & a_{n-1} & a_{n-2} & a_{n-3} & . \end{vmatrix} .$$

(iii) If $\log\log(1+x) = \log x - b_1 x - b_2 x^2 - b_3 x^3 - \dots$, show that

$$nb_n = \begin{vmatrix} \frac{1}{2} & 1 & 0 & 0 & . \\ \frac{2}{3} & \frac{1}{2} & 1 & 0 & . \\ . & . & . & . & . \\ \frac{n-1}{n} & \frac{1}{n-1} & \frac{1}{n-2} & \frac{1}{n-3} & . \\ \frac{n}{n+1} & \frac{1}{n} & \frac{1}{n-1} & \frac{1}{n-2} & . \end{vmatrix} .$$

6. (i) If $\exp(a_1 x - a_2 x^2 + a_3 x^3 - \dots) = 1 + b_1 x + \frac{b_2 x^2}{2!} + \frac{b_3 x^3}{3!} + \dots$, show that

$$b_n = \begin{vmatrix} a_1 & 1 & 0 & 0 & . \\ 2a_2 & a_1 & 2 & 0 & . \\ 3a_3 & 2a_2 & a_1 & 3 & . \\ . & . & . & . & . \\ na_n & (n-1)a_{n-1} & (n-2)a_{n-2} & (n-3)a_{n-3} & . \end{vmatrix} .$$

(ii) If $e^{-1}(1+x)^{1/x} = 1 - b_1 x + \frac{b_2 x^2}{2!} - \frac{b_3 x^3}{3!} + \dots$, show that

$$b_n = \begin{vmatrix} \frac{1}{2} & 1 & 0 & 0 & . \\ \frac{2}{3} & \frac{1}{2} & 2 & 0 & . \\ \frac{3}{4} & \frac{2}{3} & \frac{1}{2} & 3 & . \\ . & . & . & . & . \\ \frac{n}{n+1} & \frac{n-1}{n} & \frac{n-2}{n-1} & \frac{n-3}{n-2} & . \end{vmatrix} .$$

and obtain the expansion of $$(1-x)^{1/x}$$
in similar form.

7. If ϕ denotes a polynomial function of its argument and Ω the operator

$$a_1\frac{\partial}{\partial a_0}+a_2\frac{\partial}{\partial a_1}+\ldots+a_n\frac{\partial}{\partial a_{n-1}},$$

obtain the expansion

$$\phi\left(a_0+a_1x+\frac{a_2x^2}{2!}+\ldots+\frac{a_nx^n}{n!}\right)=\phi(a_0)+x\Omega\phi(a_0)+\frac{x^2}{2!}\Omega^2\phi(a_0)+\ldots.$$

8. If

$$(1-e^x)^{-n}\equiv -B_1x^{-1}+B_2x^{-2}+\ldots(-)^nB_nx^{-n}+A_0+A_1x+A_2x^2+\ldots \text{ to } \infty,$$

show that $B_r/(r-1)!$ is equal to the sum of the products $r-1$ together of the $n-1$ reciprocals of integers

$$1,\ \tfrac{1}{2},\ \tfrac{1}{3},\ldots,\ (n-1)^{-1}.$$

9. Show that $\frac{1}{2}x+x/(e^x-1)$ is an even function of x.

If $$\frac{x}{e^x-1}=1-\frac{x}{2}+\frac{B_1x^2}{2!}-\frac{B_2x^4}{4!}+\frac{B_3x^6}{6!}-\ldots \text{ to } \infty,$$

obtain the expansions

$$\frac{x}{e^x+1}=\frac{x}{2}-\frac{(2^2-1)B_1x^2}{2!}+\frac{(2^4-1)B_2x^4}{4!}-\frac{(2^6-1)B_3x^6}{6!}+\ldots,$$

$$\tan x=\frac{2^2(2^2-1)B_1x}{2!}+\frac{2^4(2^4-1)B_2x^3}{4!}+\frac{2^6(2^6-1)B_3x^5}{6!}+\ldots,$$

$$\cot x=\frac{1}{x}-\frac{2^2B_1x}{2!}-\frac{2^4B_2x^3}{4!}-\frac{2^6B_3x^5}{6!}-\ldots,$$

$$\operatorname{cosec} x=\frac{1}{x}+\frac{2(2-1)B_1x}{2!}+\frac{2(2^3-1)B_2x^3}{4!}+\frac{2(2^5-1)B_3x^5}{6!}+\ldots,$$

$$\log\left(\frac{\tan x}{x}\right)=\frac{2^3(2-1)B_1x^2}{2.2!}+\frac{2^5(2^3-1)B_2x^4}{4.4!}+\frac{2^7(2^5-1)B_3x^6}{6.6!}+\ldots,$$

$$\log\left(\frac{x}{\sin x}\right)=\frac{2^2B_1x^2}{2.2!}+\frac{2^4B_2x^4}{4.4!}+\frac{2^6B_3x^6}{6.6!}+\ldots.$$

Write down the expansions for the corresponding hyperbolic functions.

10. (i) With the notation of the previous example prove that, if r and n are positive integers,

$$1^r+2^r+\ldots+n^r$$
$$=\frac{(n+1)^{r+1}}{r+1}-\tfrac{1}{2}(n+1)^r+\frac{rB_1(n+1)^{r-1}}{2!}-\frac{r(r-1)(r-2)B_2(n+1)^{r-3}}{4!}+\ldots.$$

With what term does the series end?

(ii) Assuming that $$\frac{x}{e^x-1}=1-\tfrac{1}{2}x+\sum_{n=1}^{\infty}\frac{2x^2}{x^2+4n^2\pi^2},$$

obtain the formula $$B_r=(2r)!\,2^{1-2r}\pi^{-2r}\sum_{n=1}^{\infty}n^{-2r}.$$

(iii) Show that the coefficients of x^{2n} $(n \geqslant 1)$ in the expansions of

$$x\cot x, \qquad x\operatorname{cosec} x, \qquad x\tan x, \qquad \sec x$$

are the respective infinite series

$$-\frac{2}{\pi^{2n}}\left\{\frac{1}{1^{2n}}+\frac{1}{2^{2n}}+\frac{1}{3^{2n}}+\dots\right\}, \qquad \frac{2}{\pi^{2n}}\left\{\frac{1}{1^{2n}}-\frac{1}{2^{2n}}+\frac{1}{3^{2n}}-\dots\right\},$$

$$\frac{2^{2n+1}}{\pi^{2n}}\left\{\frac{1}{1^{2n}}+\frac{1}{3^{2n}}+\frac{1}{5^{2n}}+\dots\right\}, \qquad \frac{2^{2n+2}}{\pi^{2n+1}}\left\{\frac{1}{1^{2n+1}}-\frac{1}{3^{2n+1}}+\frac{1}{5^{2n+1}}-\dots\right\}.$$

Show that these expansions are permissible, if $|x| < \frac{1}{2}\pi$.

11. With the notation of example 8 obtain, by division of series or otherwise, the formulae

$$\frac{2^{2r}B_r}{(2r)!} = \begin{vmatrix} \frac{2}{3!} & 1 & 0 & . \\ \frac{4}{5!} & \frac{1}{3!} & 1 & . \\ . & . & . & . \\ \frac{2r}{(2r+1)!} & \frac{1}{(2r-1)!} & \frac{1}{(2r-3)!} & . \end{vmatrix},$$

$$\frac{(2^{2r}-2)B_r}{(2r)!} = \begin{vmatrix} \frac{1}{3!} & 1 & 0 & . \\ \frac{1}{5!} & \frac{1}{3!} & 1 & . \\ . & . & . & . \\ \frac{1}{(2r+1)!} & \frac{1}{(2r-1)!} & \frac{1}{(2r-3)!} & . \end{vmatrix},$$

$$\frac{2^{2r}(2^{2r}-1)B_r}{(2r)!} = \begin{vmatrix} 1 & 1 & 0 & . \\ \frac{1}{3!} & \frac{1}{2!} & 1 & . \\ . & . & . & . \\ \frac{1}{(2r-1)!} & \frac{1}{(2r-2)!} & \frac{1}{(2r-4)!} & . \end{vmatrix}.$$

12. Expand in ascending powers of x

(i) $\sin(m\sinh^{-1}x+\alpha)$, (ii) $\sinh(m\sin^{-1}x+\alpha)$,

(iii) $\sinh(m\sinh^{-1}x+\alpha)$,

and discuss the convergence of the expansions.

Obtain the infinite series

$$\sinh\tfrac{1}{2}m\pi = m+\frac{m(m^2+1^2)}{3!}+\frac{m(m^2+1^2)(m^2+3^2)}{5!}+\dots,$$

$$\cosh\tfrac{1}{2}m\pi = 1+\frac{m^2}{2!}+\frac{m^2(m^2+2^2)}{4!}+\dots,$$

$$\sinh\tfrac{1}{2}m\pi = \tfrac{1}{2}m\pi\left\{1+\frac{m^2}{2^2}+\frac{m^2(m^2+2^2)}{2^2.4^2}+\dots\right\},$$

$$\cosh\tfrac{1}{2}m\pi = 1+\frac{m^2}{1^2}+\frac{m^2(m^2+1^2)}{1^2.3^2}+\frac{m^2(m^2+1^2)(m^2+3^2)}{1^2.3^2.5^2}+\dots.$$

If $\lambda = \log(\sqrt{2}+1)$, obtain the infinite series

$$\sin m\lambda = m - \frac{m(m^2+1^2)}{3!} + \frac{m(m^2+1^2)(m^2+3^2)}{5!} - \dots,$$

$$\cos m\lambda = 1 - \frac{m^2}{2!} + \frac{m^2(m^2+2^2)}{4!} - \dots,$$

$$\sinh m\lambda = m + \frac{m(m^2-1^2)}{3!} + \frac{m(m^2-1^2)(m^2-3^2)}{5!} + \dots,$$

$$\cosh m\lambda = 1 + \frac{m^2}{2!} + \frac{m^2(m^2-2^2)}{4!} + \dots.$$

13. Expand $\qquad \exp(m\sin^{-1}x), \qquad \exp(m\sinh^{-1}x)$

in ascending powers of x, and deduce the corresponding expansions of

$$(\sin^{-1}x)^n, \qquad (\sinh^{-1}x)^n,$$

where n is a positive integer

14. Expand $\{x+\sqrt{(x^2+a)}\}^m$ in descending powers of x, and show that, if $\theta > \log(\sqrt{2}+1)$,

$$e^{m\theta} = (2\sinh\theta)^m - m(2\sinh\theta)^{m-2} + \frac{m(m-3)}{2!}(2\sinh\theta)^{m-4} + \dots \text{ to } \infty.$$

Obtain the expansions

$$x^m = (x+x^{-1})^m - m(x+x^{-1})^{m-2} + \frac{m(m-3)}{2!}(x+x^{-1})^{m-4} + \dots \text{ to } \infty \qquad (x > 1),$$

$$= (x-x^{-1})^m + m(x-x^{-1})^{m-2} + \frac{m(m-3)}{2!}(x-x^{-1})^{m-4} + \dots \text{ to } \infty \qquad (x > \sqrt{2}+1).$$

15. Expand in ascending powers of x

$$\frac{\sinh(m\sin^{-1}x+\alpha)}{\sqrt{(1-x^2)}}, \qquad \frac{\sin(m\sinh^{-1}x+\alpha)}{\sqrt{(1-x^2)}}, \qquad \frac{\sinh(m\sinh^{-1}x-\alpha)}{\sqrt{(1+x^2)}},$$

$$\frac{(\sin^{-1}x)^n}{\sqrt{(1-x^2)}}, \qquad \frac{(\sinh^{-1}x)^n}{\sqrt{(1+x^2)}},$$

where n is a positive integer

16. If $|\alpha| < \frac{1}{4}\pi$, prove that

(i) $\sin\alpha + \frac{1}{3}\sin^3\alpha + \frac{1}{5}\sin^5\alpha + \dots$ to ∞

$$= \tan\alpha - \frac{1}{2}\frac{\tan^3\alpha}{3} + \frac{1.3}{2.4}\frac{\tan^5\alpha}{5} - \dots \text{ to } \infty,$$

(ii) $\frac{\sin^2\alpha}{2} + (1+\frac{1}{3})\frac{\sin^4\alpha}{4} + (1+\frac{1}{3}+\frac{1}{5})\frac{\sin^6\alpha}{6} + \dots$ to ∞

$$= \frac{\tan^2\alpha}{2} - \frac{2}{3}\frac{\tan^4\alpha}{4} + \frac{2.4}{3.5}\frac{\tan^6\alpha}{6} - \dots \text{ to } \infty;$$

and, if $|\beta| < \log(\sqrt{2}+1)$, prove that

(iii) $\sinh\beta - \frac{1}{3}\sinh^3\beta + \frac{1}{5}\sinh^5\beta - \dots$ to ∞

$$= \tanh\beta + \frac{1}{2}\frac{\tanh^3\beta}{3} + \frac{1.3}{2.4}\frac{\tanh^5\beta}{5} + \dots \text{ to } \infty,$$

(iv) $\frac{\sinh^2\beta}{2} - (1+\frac{1}{3})\frac{\sinh^4\beta}{4} + (1+\frac{1}{3}+\frac{1}{5})\frac{\sinh^6\beta}{6} - \dots$ to ∞

$$= \frac{\tanh^2\beta}{2} + \frac{2}{3}\frac{\tanh^4\beta}{4} + \frac{2.4}{3.5}\frac{\tanh^6\beta}{6} + \dots \text{ to } \infty.$$

17. Show that in the closed interval $(-\pi, \pi)$

$$2\sum_{n=1}^{\infty}(-)^{n-1}\frac{\cos nx}{n^{2r}} = \frac{(2^{2r}-2)\pi^{2r}B_r}{(2r)!} - \frac{(2^{2r-2}-2)\pi^{2r-2}B_{r-1}x^2}{(2r-2)!\,2!} + \\ + \ldots(-)^{r-1}\frac{(2^2-2)\pi^2 B_1 x^{2r-2}}{2!(2r-2)!} + (-)^r\frac{x^{2r}}{(2r)!},$$

and in the closed interval $(0, 2\pi)$

$$2\sum_{n=1}^{\infty}\frac{\cos nx}{n^{2r}} = \frac{(2\pi)^{2r}B_r}{(2r)!} - \frac{(2\pi)^{2r-2}B_{r-1}}{(2r-2)!\,2!} + \\ + \ldots(-)^{r-1}\frac{(2\pi)^2 B_1 x^{2r-2}}{2!(2r-2)!} + (-)^r\frac{\pi x^{2r-1}}{(2r-1)!} - (-)^r\frac{x^{2r}}{(2r)!},$$

where B_1, B_2,... are Bernoulli's numbers as defined in example 9.

18. If r is a positive integer and

$$f(r, x) \equiv 1 - \frac{x^2}{(2r+1)(2r+2)} + \frac{x^4}{(2r+1)(2r+2)(2r+3)(2r+4)} - \ldots \text{ to } \infty,$$

show that, if $0 < x \leqslant 2\pi$,

$$\sum_{n=1}^{\infty} f(r, nx) = \frac{\pi}{4rx} - \frac{1}{2},$$

and, if $0 < |x| \leqslant \pi$,

$$\sum_{n=1}^{\infty}(-)^{n-1}f(r, nx) = \frac{1}{2}.$$

19. Show that in the closed interval $(-\pi, \pi)$

$$\sum_{n=1}^{\infty}(-)^{n-1}\frac{\sin nx}{n(n^2+p^2)} = \frac{x - \pi \operatorname{cosech} p\pi \sinh px}{2p^2},$$

$$\sum_{n=1}^{\infty}(-)^{n-1}\frac{\cos nx}{n^2+p^2} = \frac{1 - p\pi \operatorname{cosech} p\pi \cosh px}{2p^2},$$

and that

$$\pi \coth p\pi = \frac{1}{p} + \sum_{n=1}^{\infty}\frac{2p}{p^2+n^2}, \qquad \pi \operatorname{cosech} p\pi = \frac{1}{p} + \sum_{n=1}^{\infty}(-)^n\frac{2p}{p^2+n^2}.$$

Deduce that

$$\sinh\theta = \theta\prod_{n=1}^{\infty}\left(1 + \frac{\theta^2}{n^2\pi^2}\right), \qquad \cosh\theta = \prod_{n=1}^{\infty}\left(1 + \frac{\theta^2}{(n+\frac{1}{2})^2\pi^2}\right).$$

20. Show that

$$1 - \frac{1}{3^3} + \frac{1}{5^3} - \ldots \text{ to } \infty = \frac{\pi^3}{32},$$

$$1 - \frac{1}{3^5} + \frac{1}{5^5} - \ldots \text{ to } \infty = \frac{5\pi^5}{1536},$$

$$\frac{1}{1(1^2-p^2)} - \frac{1}{3(3^2-p^2)} + \frac{1}{5(5^2-p^2)} - \ldots \text{ to } \infty = \frac{\pi(\sec\frac{1}{2}p\pi - 1)}{4p^2},$$

$$\frac{1}{1(1^2+p^2)} - \frac{1}{3(3^2+p^2)} + \frac{1}{5(5^2+p^2)} - \ldots \text{ to } \infty = \frac{\pi(1 - \operatorname{sech}\frac{1}{2}p\pi)}{4p^2}.$$

21. Show that in the closed interval $(-\pi, \pi)$

$$\sum_{n=1}^{\infty} \frac{(-)^{n-1} n \sin nx}{\{(n-p)^2+q^2\}\{(n+p)^2+q^2\}}$$
$$= \frac{\pi(\sin p\pi \cosh q\pi \cos px \sinh qx - \cos p\pi \sinh q\pi \sin px \cosh qx)}{4pq(\sin^2 p\pi + \sinh^2 q\pi)}$$

$$\sum_{n=1}^{\infty} \frac{(-)^{n-1}\{(p^2+q^2)^2 - n^2(p^2-q^2)\}\sin nx}{n\{(n-p)^2+q^2\}\{(n+p)^2+q^2\}}$$
$$= \tfrac{1}{2}x - \frac{\pi(\sin p\pi \cosh q\pi \sin px \cosh qx + \cos p\pi \sinh q\pi \cos px \sinh qx)}{2(\sin^2 p\pi + \sinh^2 q\pi)}$$

and that

$$\frac{\tfrac{1}{2}\pi \sin 2p\pi}{\sin^2 p\pi + \sinh^2 q\pi} = \frac{p}{p^2+q^2} + 2p\sum_{n=1}^{\infty} \frac{p^2-q^2-n^2}{\{(n-p)^2+q^2\}\{(n+p)^2+q^2\}},$$

$$\frac{\tfrac{1}{2}\pi \sinh 2q\pi}{\sin^2 p\pi + \sinh^2 q\pi} = \frac{q}{p^2+q^2} + 2q\sum_{n=1}^{\infty} \frac{p^2+q^2+n^2}{\{(n-p)^2+q^2\}\{(n+p)^2+q^2\}},$$

$$\frac{\pi \sin p\pi \cosh q\pi}{\sin^2 p\pi + \sinh^2 q\pi} = \frac{p}{p^2+q^2} + 2p\sum_{n=1}^{\infty} \frac{(-)^n(p^2+q^2-n^2)}{\{(n-p)^2+q^2\}\{(n+p)^2+q^2\}},$$

$$\frac{\pi \cos p\pi \sinh q\pi}{\sin^2 p\pi + \sinh^2 q\pi} = \frac{q}{p^2+q^2} + 2q\sum_{n=1}^{\infty} \frac{(-)^n(p^2+q^2+n^2)}{\{(n-p)^2+q^2\}\{(n+p)^2+q^2\}}.$$

Deduce that

$$\sin^2 p\pi + \sinh^2 q\pi = \pi^2(p^2+q^2)\prod_{n=1}^{\infty}\left\{\left(1-\frac{p}{n}\right)^2 + \frac{q^2}{n^2}\right\}\left\{\left(1+\frac{p}{n}\right)^2 + \frac{q^2}{n^2}\right\},$$

$$\tan^{-1}\left(\frac{\tan p\pi}{\tanh q\pi}\right) = \tan^{-1}\frac{p}{q} + \sum_{n=1}^{\infty} \tan^{-1}\frac{2pq}{n^2-p^2+q^2}.$$

22. Show that

$$\sum_{n=1}^{\infty} \frac{\cos^n \alpha \sin n\alpha}{n} = \tfrac{1}{2}\pi - \alpha \qquad (0 < \alpha < \pi),$$

$$\sum_{n=1}^{\infty} \frac{\cos^n \alpha \cos n\alpha}{n} = \log \operatorname{cosec} \alpha \qquad (0 < \alpha < \pi).$$

Sum the infinite series

$$\tfrac{1}{2}\sin\alpha + \tfrac{1}{3}\sin 2\alpha + \tfrac{1}{4}\sin 3\alpha + \ldots, \qquad \sin 2\alpha + \tfrac{1}{2}\sin 3\alpha + \tfrac{1}{3}\sin 4\alpha + \ldots,$$

$$\sum_{n=1}^{\infty}\left\{\frac{\sin(p+1)\alpha}{\sin p\alpha}\right\}^n \frac{\cos(n\alpha+\beta)}{n}, \qquad \sum_{n=1}^{\infty}\left\{\frac{\cos(p+1)\alpha}{\cos p\alpha}\right\}^n \frac{\cos(n\alpha+\beta)}{n},$$

stating any necessary limitations on α.

23. If m is an integer greater than 2 and p is any real number, show that the pth powers of the roots of the algebraic equation

$$xy^m - y + 1 = 0$$

satisfy the hypergeometric differential equation

$$\delta \prod_{s=1}^{m-1} \left(\delta - 1 + \frac{p+s}{m-1}\right) y^p = \frac{m^m x}{(m-1)^{m-1}} \prod_{s=0}^{m-1} \left(\delta + \frac{p+s}{m}\right) y^p,$$

and express each of these pth powers in hypergeometric series, distinguishing the cases

$$x \gtrless \frac{(m-1)^{m-1}}{m^m}.$$

24. If m, n are integers such that

$$m > 2, \qquad m > n > 0,$$

and p is any real number, show that the pth powers of the roots of the algebraic equation

$$xy^m - y^n + 1 = 0$$

satisfy the hypergeometric differential equation of order m

$$\prod_{s=0}^{n-1} (\delta - s) \prod_{s=1}^{m-n} \left(\delta - n + \frac{p+ns}{m-n}\right) y^p = \frac{m^m x^n}{n^n (m-n)^{m-n}} \prod_{s=0}^{m-1} \left(\delta + \frac{p+ns}{m}\right) y^p,$$

and express each of these pth powers in hypergeometric series, distinguishing the cases

$$x^n \gtrless \frac{n^n (m-n)^{m-n}}{m^m}.$$

25. If m, n are integers such that

$$m > 2, \qquad m > n > 0,$$

show that the logarithms of the roots of the algebraic equation

$$xy^m - y^n + 1 = 0$$

satisfy the hypergeometric equation of order $m+1$

$$\delta^2 \prod_{s=1}^{n} (\delta - s) \prod_{s=1}^{m-n-1} \left(\delta - n + \frac{ns}{m-n}\right) \log y = \frac{m^n x^n}{n^n (m-n)^{m-n}} \delta^2 \prod_{s=1}^{m-1} \left(\delta + \frac{ns}{m}\right) \log y,$$

and express each of these logarithms in hypergeometric series, distinguishing the cases

$$x^n \gtrless \frac{n^n (m-n)^{m-n}}{m^m}.$$

INDEX

The numbers refer to pages. Italicized references are to definitions only.

NOTATION

PRINTED IN
GREAT BRITAIN
AT THE
UNIVERSITY PRESS
OXFORD
BY
JOHN JOHNSON
PRINTER
TO THE
UNIVERSITY